Subject: Re: Choosing the best way to go full UNICODE
Posted by cbpporter on Wed, 14 Jun 2017 21:31:09 GMT
View Forum Message <> Reply to Message

PS: that is a special composition. I went over the data over and over again and I found no good reason to handle box decomposition.

It is not like U++ will check to see if the font supports that character and if not, decompose it and build CJK in a small box on the fly.

Decompositions that start with <smth> are all special, like <font>, meaning that you can decompose that character if you are doing font substitution to an approximation or <square>, meaning that the code point is multiple characters arranged in a square or like <fraction>, when you have 1/2 as a single code point and you can decompose it as 1/2, using 3 code points.

I choose to ignore all these for now since I can't figure out how to offer any worthwhile feature related to these special substitutions. I don't even need normal decompositions, but it is pretty cool to decompose diacritics and replace some bits since I'm an European and my native language uses diacritics.

As for NCF and NDF I only found two good use cases: string equality and search. With the forms, you don't compare code points, but glyphs, without building glyphs. If two string look the same on your display, but have different code points due to diacritics, it is very useful to tell if they are

encoded as a "t" with a composition mark to be identified as the same string.