

Deprecation of 3 Tibetan Characters

Author: Élie Roux, lead developer at Buddhist Digital Resource Center (<https://www.tbrc.org/>)

Date of Submission: March 23rd 2018

This is a proposal to deprecate the following characters

- U+0F00 ཨོ TIBETAN SYLLABLE OM
- U+0F02 འུྱ TIBETAN MARK GTER YIG MGO -UM RNAM BCAD MA
- U+0F03 འུྲ TIBETAN MARK GTER YIG MGO -UM GTER TSHEG MA

Currently these three characters are not indicated as being composed, as we can see in [UnicodeData.txt](#):

```
0F00;TIBETAN SYLLABLE OM;Lo;0;L;;;;;N;;;;;
0F02;TIBETAN MARK GTER YIG MGO -UM RNAM BCAD MA;So;0;L;;;;;N;;;;;
0F03;TIBETAN MARK GTER YIG MGO -UM GTER TSHEG MA;So;0;L;;;;;N;;;;;
```

But these characters are composed as such (between parenthesis is the rendering in a popular Tibetan Font):

- 0F00 (ཨོ) ≡ 0F68 0F7C 0F7E (ཨོ)
- 0F02 (འུྱ) ≡ 0F60 0F74 0F82 0F7F (འུྱ)
- 0F03 (འུྲ) ≡ 0F60 0F74 0F82 0F14 (འུྲ)

There is an acknowledgment of the composition of 0F00 in the [DUCET](#):

```
0F00 ; [ .2F19.0020.0004 ] [ .2F30.0020.0004 ] [ .0000.00C4.0004 ]
```

but not for 0F02 or 0F03:

```
0F02 ; [ *053D.0020.0002 ]
0F03 ; [ *053E.0020.0002 ]
```

The composition is also acknowledged in [L2/08-220](#), and the composition of 0F02 and 0F03 is indicated in their name:

- -UM is for 0F74 (TIBETAN VOWEL SIGN U) 0F82 (TIBETAN SIGN NYI ZLA NAA DA, often indicated as M in transliterations)
- RNAM BCAD is for 0F7F (TIBETAN SIGN RNAM BCAD)
- GTER TSHEG is for 0F14 (TIBETAN MARK GTER TSHEG)

Because of Unicode's [normalization stability policy](#), the composition indication cannot be added. And because the characters have been in Unicode for more than two years, the [Change Management for the Unicode Collation Algorithm](#) makes any change of the weights of 0F02 and 0F03 in the DUCET very unlikely.

My understanding of the situation is that we have different incompatible ways of encoding the same thing.

My proposal is to deprecate 0F00, 0F02 and 0F03 as it seems the best possible way of fixing this situation.