A minimum set of glyph corrections needed for the 10.0 beta Mongolian code chart

Liang Hai (梁海) <lianghai@gmail.com>

12 May 2017

This document mainly examines the "type 1" issues (8 glyph issues, involving 7 characters) in the attached spreadsheet by Shen Yilei (沈逸磊).

§ To be fixed for 10.0:

1836 MONGOLIAN LETTER YA
~ 1836 180C third form (medial)

- Wrong since Unicode 7.0 in the code chart but correct in the StandardizedVariants.html file. Refer to the latest StandardizedVariants.html (Unicode 8.0) for the correct form.
- Should be identical to 1822 MONGOLIAN LETTER I ~ 1822 first form (final):

  \[ \text{\textcircled{1}} \]

- The glyph mentioned above can be used directly.

1846 MONGOLIAN LETTER TODO O
~ 1846 180B second form (medial)

- Wrong since Unicode 7.0 in the code chart but correct in the StandardizedVariants.html file. Refer to the latest StandardizedVariants.html (Unicode 8.0) for the correct form.
- Should look like 1848 MONGOLIAN LETTER TODO OE ~ 1848 180B second form (medial) without the upward tail, or a sequence of <1820 MONGOLIAN LETTER A ~ 1820 first form (medial), 1846 MONGOLIAN LETTER TODO O ~ 1846 first form (medial)>:
A correct glyph might not exist in the font (Mongolian Baiti), but can be created according to the description above.

185E MONGOLIAN LETTER SIBE I
~ 185E 180C third form (medial)

- Wrong from the beginning (Unicode 3.2) in both the code chart and the StandardizedVariants.html file.
- Should be identical to 1873 MONGOLIAN LETTER MANCHU I ~ 1873 180D fourth form (medial):

- The glyph mentioned above can be used directly.

§ Already fixed in the latest 10.0 draft:

1820 MONGOLIAN LETTER A
~ 1820 180B second form (medial)

- Wrong since Unicode 9.0. Refer to the latest StandardizedVariants.html (Unicode 8.0) for the correct form.
- Should look like a sequence of <1820 MONGOLIAN LETTER A ~ 1820 first form (medial), 1820 MONGOLIAN LETTER A ~ 1820 first form (medial)>:

- A correct glyph should already exist in the font at the expected position.

§ Typographical; not to be fixed for 10.0:
The author considers the following two cases to be typographical, therefore not urgent for 10.0 although indeed errors and should be corrected in the future (actually the glyphs of these two seem to have been somehow swapped).

**1832 MONGOLIAN LETTER TA**
- 1832 180B second form (medial)

**1868 MONGOLIAN LETTER SIBE TA**
- 1868 180B second form (initial)

And about the following two, although Shen Yilei is mostly concerned about the stroke endings (actually the stroke beginning in the first one is also an error), the author considers these are also typographical issues as much as the two above.

**186F MONGOLIAN LETTER SIBE ZA**
- 186F 180B second form (initial)
- 186F 180B second form (medial)

---

**§ Type 2, editorially swapped forms:**

It's believed that the current situation is because glyphs of these two forms were somehow editorially swapped.

**1836 MONGOLIAN LETTER YA**
- 1836 first form (medial)
- 1836 180B second form (medial)

- Wrong since Unicode 7.0 in the code chart but correct in the StandardizedVariants.html file. Refer to the latest StandardizedVariants.html (Unicode 8.0) for the correct forms.
- ~ 1836 first form (medial) should be identical to 1822 MONGOLIAN LETTER I ~ 1822 first form (medial), a hookless form:

```
/\      /
|       |
\  --  /  
```

- ~ 1836 180B second form (medial) should resemble 1836 MONGOLIAN LETTER YA ~ 1836 first form (initial), a hooked form:
The glyphs mentioned above can be used directly. Or the current two glyphs in 10.0 beta can be simply swapped to meet the description above.

§ Type 3, missing isolate forms:

These four isolate forms should’ve been in the code chart due to the exactly same reason why U+1824 MONGOLIAN LETTER U and U+1826 MONGOLIAN LETTER UE has their first form (isolate) explicitly listed.

Vowel letters of the Mongolian script have true isolate forms therefore existence of these forms are even more significant than the "type 3" isolate forms of consonant letters.

185D MONGOLIAN LETTER SIBE E
~ 185D ḡ first form (isolate)

185E MONGOLIAN LETTER SIBE I
~ 185E ḡ first form (isolate)

1861 MONGOLIAN LETTER SIBE U
~ 1861 ḡ first form (isolate)

1873 MONGOLIAN LETTER MANCHU I
~ 1873 ḡ first form (isolate)

The four needed glyphs are identical to the representative glyphs of U+1821 MONGOLIAN LETTER E, U+1822 MONGOLIAN LETTER I, U+1825 MONGOLIAN LETTER OE, U+1822 MONGOLIAN LETTER I, respectively.

§ Appendix:

List of editorial errors in the Mongolian code chart of U10beta
Additional notes from Shen Yilei:

As for Type 4 & 5: Admittedly it is important to expose the problems in the present standard for Mongolian, but it is not the correct way to postulate new variants for the undefined slots based largely on personal judgments rather than a warranted resolution. Otherwise one may ask why should only those unsanctioned variants be recognized but not some others, say, initial and final for U+180A NIRUGU. As far as I can see, the only correct way to expose the problems is to leave the present standard as it was.

As for Type 3: The hidden logic behind the U10beta code chart, in my interpretation, is that all positional variants shall be listed somewhere. For U+1820-1823, the first isolates are identical with the representative glyphs, hence reasonably omitted; but for U+1824, the representative glyph is identical with the first INITIAL rather than ISOLATE, resulting to the explicit listing of otherwise implicit first isolates. The Type 3 problems are the omissions of dedicated first isolates, which are expected to be explicitly listed in a consistent logic.

In general, the code chart is not expected to give more information beyond the agreed-upon resolution (substantially equivalent to TR170 except for a few minor editorial errors and later amendments) of the present standard for Mongolian, which constitutes a de facto amendment. In this sense Type 4 & 5 problems are obviously errors. Type 3 problems are errors in that they unexpectedly imply the absence of some first isolates, and the readers are, at best, at risk of having a reading contradicting the resolution that the omitted variants are deleted. On the contrary, Type 6 problems are by no means wrong because they are systematically left out; but it is desired for these first isolates identical with the representative glyphs to be listed explicitly, to preclude the current idiosyncrasies in implementing these forms in the industry from happening again partially owing to the unclear specification in the code chart.