

DOC TYPE: Working Group Document
 TITLE: Mongolian Ad Hoc Report
 SOURCE: Deborah Anderson, SEI, UC Berkeley
 STATUS: Expert contribution
 ACTION: For consideration by WG2
 DATE: 29 September 2016



The following experts were in attendance: Deborah Anderson, Ken Whistler, Michel Suignard, Lisa Moore, Ou Rileke, Nasun-urt, Wei Lin Mei, Wu Yingzhe, Chen Zhuang, Lu Qin, and Greg Eck (by phone).

The discussion was based on the following documents:

- N4757 Modifications to Mongolian Encoding in UCS (=L2/16-261)
- N4753 WG2 #65 Mongolian Discussion Points (=L2/16-259)
- N4752 Mongolian Base Forms, Positional Forms, & Variant Forms (=L2/16-258)
- N4763 Comments on Mongolian, Small Khitan, and other WG2 #65 document (=L2/16-266)

The following points were agreed to in the ad hoc meeting:

1. from N4753 WG2 #65 Mongolian Discussion Points (=L2/16-259)

- RE: #1 in N4753: NNBSB Deficiency
 - The problems with the proposal for encoding a new Mongolian character to replace the NNBSB were discussed at length, but no consensus was reached.
 - Greg Eck, in conjunction with Nasun-urt, will create a document that identifies the issues, with relevant data. The document should also include a comparison showing why a new Mongolian character would be preferable to NNBSB.
 - The change of general category for NNBSB from gc=Zs to gc=Cf was discussed as one potential way to handle NNBSB, and might not involve the encoding of a new character.
- RE: #2 in N4753: U+1887 Deficiency
 - The experts confirmed that fourth final form (shown below) is needed:

 - The Editor of ISO 10646 should add the sequence U+1887 + U+FE00 [VS01] fifth form final with the following glyph:

 - The ad hoc group recommends the Editor of ISO 10646 make the following changes to [StandardizedVariants.txt](#):

(Current)

1887 180B; second form; isolate final # MONGOLIAN LETTER ALI GALI A
 1887 180C; third form; final # MONGOLIAN LETTER ALI GALI A
 1887 180D; fourth form; final # MONGOLIAN LETTER ALI GALI A

(Proposed change, noted in **bold**; Note: the second form isolate deprecated)

1887 180B; second form; isolate final # MONGOLIAN LETTER ALI GALI A










1887 180C; third form; final # MONGOLIAN LETTER ALI GALI A

1887 180D; fourth form; final # MONGOLIAN LETTER ALI GALI A











1887 FE00; fifth form; final # MONGOLIAN LETTER ALI GALI A

- The UTC will need to discuss how to handle the sequence of the second form isolate, since the glyph for the newly proposed sequence for the fifth form final is identical to the second form isolate.
- RE: #3 in N4752: Unicode 9.0 Corrections
 - The ad hoc group recommends that the Editor of ISO 10646 make changes to U+182D as follows:

(Current)

182D  MONGOLIAN LETTER GA
→ 0433 Г cyrillic small letter ghe
~ 182D  first form (initial)
~ 182D  first form (medial)
~ 182D  first form (final)
~ 182D 180B  second form (initial)
~ 182D 180B  second form (medial)
~ 182D 180B  feminine form (final)
~ 182D 180C  third form (medial)
~ 182D 180D  feminine form (medial)

(Proposed change)

182D  MONGOLIAN LETTER GA
→ 0433 Г cyrillic small letter ghe
~ 182D  first form (initial)
~ 182D  first form (medial)
~ 182D  first form (final)
~ 182D 180B  second form (initial)
~ 182D 180B  second form (medial)
~ 182D 180B  second form (final)
~ 182D 180C  third form (medial)
~ 182D 180C  feminine third form (final)
~ 182D 180D  feminine form (medial)

- The ad hoc group recommends that the Editor of ISO 10646 make the following changes to [StandardizedVariants.txt](#):

(Current)

182D 180B; second form; initial medial # MONGOLIAN LETTER GA

182D 180B; feminine form; final # MONGOLIAN LETTER GA

182D 180C; third form; medial # MONGOLIAN LETTER GA

(Proposed change, noted in **bold**)

182D 180B; second form; initial medial **final** # MONGOLIAN LETTER GA

182D 180C; third form; medial # MONGOLIAN LETTER GA

182D 180C; feminine third form; final # MONGOLIAN LETTER GA

- o The ad hoc group also recommends that the Editor of ISO 10646 make the following changes:
 - Remap the font to get the 182D 180B second form final context mapped to the “z” shape.
 - Add a mapping for 182D 180C feminine third form final context to the one currently shown for the 182D 180B feminine form final context.
- o The ad hoc group recommends the Editor of ISO 10646 work with the authors of N4752 Mongolian Base Forms, Positional Forms, & Variant Forms (=L2/16-258 = DS01) to identify other letters whose feminine forms will require fixes.
- RE: #4 Nomenclature Standardization
 - o The ad hoc group recommends that the authors create an independent publication in their preferred format, since the current chart generation process must be able to handle over 121,000 characters, and cannot be re-tooled specifically for Mongolian. Such a publication could be a [Unicode Technical Note](#), or a web publication, such as Richard Ishida’s [Notes on Mongolian Variant Forms](#)
 - o Experts in the ad hoc group felt the removal of the identification of “feminine” forms from specific standardized sequences was not advisable.

2. From N4757 Modifications to Mongolian Encoding in UCS (=L2/16-261)

Part One

- RE: #1 New characters (page 2) in N4757
 - o The ad hoc group recommended the authors submit a proposal for any new characters or sequences that are not in the 5th edition of ISO 10646 or DS01. The proposal should provide the glyph and clearly identify the base character (by code point), FVS, the context (second final form, etc.), and whether it would be an extension to DS01. Proposals with all the required information would be eligible for inclusion in PDAM 1.2 if received by the Editor of ISO 10646 in four weeks (27 October 2016).
- RE: #2 (Free) variation selectors (page 2) in N4757
 - o The ad hoc group notes that sixteen Variation Selectors are available for use by any script (see <http://www.unicode.org/charts/PDF/UFE00.pdf>), so no new variation selectors need be encoded specifically in the Mongolian block to handle this requirement.
- RE: #4 Using variation selector characters within words (page 2) in N4757
 - o The ad hoc group recommends the Editor of ISO 10646 show appropriate positional forms, for all Mongolian characters which have them, in the names list.

- RE:#6 Contents to be added (page 7)
 - The ad hoc group recommends the authors submit a proposal for any new sequences that are not in the 5th edition or DS01. (For details on what such a proposal should contain, see #1 above.)

Part Two

- RE: #1 Partial adjustments to UCS
 - The ad hoc group encouraged the authors to document which sequences with FVSes are used in mandatory ligatures. Such information is needed by implementers. An appropriate publication vehicle would be a Unicode Technical Note or other publication.
- RE: #5 Questions of groupings by scripts (page 9)
 - The ad hoc group recommended subsets of characters (by language) be included in a separate publication, such as a Unicode Technical Note, a W3C document, or other publication. The Editor of ISO 10646 can provide the data for such a subset, although it cannot be published in the standard itself.

3. Other comments

- The ad hoc group, which included Mongolian experts, agreed to collaborate in reviewing DS01 and to send comments to Greg Eck and the Editor of ISO 10646.