Insufficiency of the existing description of variant selectors and variant sequences in Mongolian Unicode character table

by Enkhdalai B.
ULAANBAATAR - 2019
Contents

• Problems on insufficiency of variation description

• Invisible Format controls are visible by default
• Problems on insufficiency of variation description
The description of variant selectors and variant sequences of Mongolian characters can be divided into three categories* depending on the complexity of the character variations.

1. Basic Simple Variations
2. Basic simple variations with additional variations
3. Complicated variations
1. **Basic Simple Variations**

182 E MONGOLIAN LETTER MA
→ 043 C м cyrillic small letter em
~ 182 E ‡ first form (initial)
~ 182 E ‡ first form (medial)
~ 182 E ‡ first form (final)

Existing description of variant selectors and variant sequences has enough information.
2. **Basic Simple Variations with additional variations**

1830 ᠊蒙古字母 SA  → 0441 c cyrillic small letter es
~ 1830 ᠊第一形 (initial)
~ 1830 ᠊第二形 (medial)
~ 1830 ᠊第三形 (final)
~ 1830 180B  second form (final)
~ 1830 180C  third form (final)

Existing description of variant selectors and variant sequences has enough information.
3. **Complicated variations**

182D 긯 MONGOLIAN LETTER GA
→ 0433 razione small letter ghe
~ 182D 긯 first form (initial)
~ 182D 긲 first form (medial)
~ 182D 긳 first form (final)
~ 182D 긲 second form (initial)
~ 182D 긲 second form (medial)
~ 182D 긲 second form (final)
~ 182D 긲 third form (medial)
~ 182D 긲 fourth form (medial)

Existing description of variant selectors and variant sequences has not fully enough information.
### Complicated variations

Initial position – default 2 variations

<table>
<thead>
<tr>
<th>голүү (gölüge)</th>
<th>гарааг (γaray)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="gölüge" /></td>
<td><img src="image2" alt="γaray" /></td>
</tr>
</tbody>
</table>
But,

In the existing description of variant sequences, there is showing only one variation of initial letter:
Thus,

Suggested modification is as follows, where is described all of two initial first forms instead of one is.

≈ 182D ❄ first form (initial)
≈ 182D ❋ first form (initial)
3. **Complicated variations**

Medial position – default 3 variations

<table>
<thead>
<tr>
<th>тагар (tayar)</th>
<th>хагда (qayda)</th>
<th>хэгэр (keger)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="%E8%92%99%E5%8F%A4%E8%AA%9E.png" alt="tayar" /></td>
<td><img src="%E8%92%99%E5%8F%A4%E8%AA%9E.png" alt="qayda" /></td>
<td><img src="%E8%92%99%E5%8F%A4%E8%AA%9E.png" alt="keger" /></td>
</tr>
</tbody>
</table>
But,

In the existing description of variant sequences, there is showing only one variation of medial letter:
Thus,

Suggested modification is as follows, where is described all of three medial first forms instead of one is.

~ 182D first form (medial)
~ 182D first form (medial)
~ 182D first form (medial)
So, modified description for U+182D «Г» will be as:

<table>
<thead>
<tr>
<th>182D Г MONGOLIAN LETTER GA</th>
<th>182D Г MONGOLIAN LETTER GA</th>
</tr>
</thead>
<tbody>
<tr>
<td>→ 0433 г cyrillic small letter ghe</td>
<td>→ 0433 г cyrillic small letter ghe</td>
</tr>
<tr>
<td>~ 182D Г first form (initial)</td>
<td>• Context driven,</td>
</tr>
<tr>
<td>~ 182D Г first form (medial)</td>
<td>~ 182D Г first form (initial)</td>
</tr>
<tr>
<td>~ 182D Г first form (final)</td>
<td>~ 182D Г first form (initial)</td>
</tr>
<tr>
<td>~ 182D 180B Г second form (initial)</td>
<td>• Context driven,</td>
</tr>
<tr>
<td>~ 182D 180B Г second form (medial)</td>
<td>~ 182D Г first form (medial)</td>
</tr>
<tr>
<td>~ 182D 180B Г second form (final)</td>
<td>~ 182D Г first form (medial)</td>
</tr>
<tr>
<td>~ 182D 180C Г third form (medial)</td>
<td>• Context driven,</td>
</tr>
<tr>
<td>~ 182D 180D Г fourth form (medial)</td>
<td>~ 182D Г first form (final)</td>
</tr>
<tr>
<td>~ 182D 180B Г second form (final)</td>
<td>• Positional variations selected by FVS</td>
</tr>
<tr>
<td>~ 182D 180B Г second form (initial)</td>
<td>~ 182D 180B Г second form (medial)</td>
</tr>
<tr>
<td>~ 182D 180B Г second form (medial)</td>
<td>~ 182D Г first form (final)</td>
</tr>
<tr>
<td>~ 182D 180B Г second form (final)</td>
<td>~ 182D Г first form (final)</td>
</tr>
<tr>
<td>~ 182D 180C Г third form (medial)</td>
<td>~ 182D 180C Г third form (medial)</td>
</tr>
<tr>
<td>~ 182D 180D Г fourth form (medial)</td>
<td>~ 182D 180D Г fourth form (medial)</td>
</tr>
</tbody>
</table>
Finally,

Complete modified description of variant selectors and variant sequences in Mongolian Unicode character table is as:
• Invisible Format controls are visible by default
Brief:

Currently all of Format controls as FVS1, FVS2 and FVS3 of Mongolian Unicode range are always invisible.

• Nobody knows is there had typed some format control character within the text context.
• Even in the visibly empty document there might be thousands of format controls are inserted.
Solution:

All of Format controls as FVS1, FVS2 and FVS3 should be:

• Default - always visible.

• Should be only single FVS used, no double and so on

• If the FVS has control over the pre-character, only then it should be invisible.
<table>
<thead>
<tr>
<th>Standard Font (OpenType)</th>
<th>Custom Font (OpenType)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FV: S3</td>
</tr>
<tr>
<td></td>
<td>FV: S2</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td>Looks like empty</td>
<td>Actually it is not empty</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Standard Font (OpenType)</th>
<th>Custom Font (OpenType)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
</tbody>
</table>

Looks like empty
Actually it is not empty

<table>
<thead>
<tr>
<th>Standard Font (OpenType)</th>
<th>Custom Font (OpenType)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
<tr>
<td></td>
<td>FV: S1</td>
</tr>
</tbody>
</table>

Looks like clean letters only
Actually it has unnecessary FVS-es
THANKS FOR YOUR ATTENTION!