## UNICODE FORUM PROPOSED CHANGES – MONGOLIAN BLOCK January 14, 2015

#### **SUMMARY**

### New FVS Character

180F-FVS4

## New FVS assignments

| 180A+FVS1-Final  | 1826+FVS2-Final   |
|------------------|-------------------|
| 1822+FVS2-Medial | 1828+FVS4-Medial  |
| 1822+FVS3-Medial | 1829+FVS1-Isolate |
| 1822+FVS1-Final  | 182C+FVS2-Initial |
| 1824+FVS1-Final  | 182D+FVS2-Initial |
| 1824+FVS2-Final  | 1843+FVS1-Final   |
| 1826+FVS4-Medial | 1887+FVS4-Final   |

## Differences with the Chinese Standard

182C+FVS2/FVS3-Medial are switched

182D+FVS1-Isolate (Chinese Standard uses a ZWJ)

## **Character Feature Changes**

1885 Baluda – Feature change from "letter" to "diacritic mark" 1886 Triple Baluda – Feature change from "letter" to "diacritic mark"

<u>Unicode Isolate Forms without FVS assignments</u> – This is an FYI at present. The issue is that there is no easy way for the user to type these code-point's Unicode form without having intrinsic understanding of the block; user will have to use a ZWJ combination; user will have to know the position of the particular code-point that is the Unicode form. It is felt that this should be a simple task to write any given Unicode form whereas in reality it is not.

| 184A (Final)         | 1862 (Final)  |
|----------------------|---------------|
| 184B (Final)         | 1863 (Final)  |
| 184F (Final)         | 1873 (Final)  |
| 1859 (Final)         | 1874 (Medial) |
| 185D (Final)         | 1875 (Final)  |
| 185E (Second Medial) | 1876 (Medial) |
| 1861 (Medial)        |               |

<u>FVS Mis-Match</u> – This is probably our hardest issue the forum has faced. There exists one post-NNBSP context where the Chinese Standard does not agree with the glyph position as determined by common font rendering engines such as the Universal Shaping Engine of Microsoft nor by Harfbuzz of Google. There also exist 6 pre-MVS contexts with the same issue. This creates confusion in the minds of the font developer. It creates further confusion in the correct placement of the FVS which is designed to simulate a given context. The issue also creates un-necessary redundancy in the probable implementation of two FVSs for one context. The suggested solution is to specify the glyph position of these 7 different letters according to the actual position as derived by the font rendering engines. See Appendix Chart – Mongolian FVS MisMatch for more details.

|                    | Chinese Standard Stated Position | Uniscribe/Harfbuzz Derived Position |
|--------------------|----------------------------------|-------------------------------------|
| U+1820(post-NNBSP) | Medial                           | Initial                             |

| U+1828(pre-MVS)   | Medial | Final |
|-------------------|--------|-------|
| U+182C(pre-MVS)#1 | Medial | Final |
| U+182C(pre-MVS)#2 | Medial | Final |
| U+182D(pre-MVS)   | Medial | Final |
| U+1835(pre-MVS)   | Medial | Final |
| U+1836(pre-MVS)   | Medial | Final |

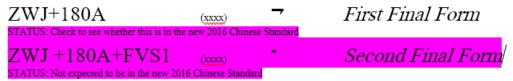
#### NEW CONTROL CHARACTER

1.) New FVS4 Control Character – There are three instances now where the FVS4 may be needed, one of which is definitely confirmed; <U+1887-Isolate><FVS1> is actually a final as confirmed by Professor Quejingzhabu.

180F (XXXX) MONGOLIAN FREE VARIATION SELECTOR FOUR

#### **NEW VARIANTS**

1.) New FVS1 assignment at Final U+180A-Nirugu – Professor Quejingzhabu's Specification 9.0 calls for a final with a hook. Current forum members are calling for a final with no hook. As the hook is more standard than no hook, it is recommended that the default be the hook and the no-hook be the variant with FVS.



2.) New FVS2 assignment at <U+1822-Medial><U+180C> Three out of six font developers implement this double-tooth U+1822 form.

3.) New FVS3 assignment at <U+1822-Medial><U+180D> This variant is needed to over-ride default context of the medial I in words such as NAIMA "eight".

4.) New FVS1 assignment at <U+1822-Final><U+180B> This 1822 final variant has been attested by various independent sources.

5.) New FVS1 assignment at <U+1824-Final><U+180B> This variant is needed for foreign spellings such as "Nassau".



6.) New FVS2 assignment at <U+1824-Final><U+180C> This variant has been attested by several independent sources as a variant of the dotted genetive case U/UE after the sequence <U+1828-N><NNBSP>.

# 

7.) New FVS3 assignment at <U+1826-Medial><U+180D> This variant is needed to over-ride the context so that we can use the DUEGER with or without the NNBSP.

8.) New FVS2 assignment at <U+1826-Final><U+180C> This variant has been attested by several independent sources as a variant of the dotted genetive case U/UE after the sequence <U+1828-N><NNBSP>.

9.) New FVS4 assignment at <U+1828-Medial><U+180F> This default form is needed with FVS assignment to over-ride native context for such rare cases as ANAR.

ZWJ+1828+FVS4+ZWJ Fifth Medial Form
STATUS: Not expected to be in the new 2016 Chinese Standard

(AN+FVS4+AR must remove the dot)

10.) New FVS1 assignment at <U+1829-Isolate><U+180B> This isolate is evidenced in the dictionary "mongyol nanggiyad üsüg-ün toli bičig" (蒙漢字典) published in 1928. This is shown in context at http://www.cneas.tohoku.ac.jp/staff/hkuri/articles/a62houkoku14.pdf



11.) New FVS2 assignment at <U+182C-Initial><U+180C> This variant is needed for the special case of QQIR.

12.) New FVS3 assignment at <U+182D><U+180C> This variant is needed for foreign loan words such as GRAM.



13.) New FVS4 assignment at <U+182D><U+180F> This default form is needed with FVS assignment to over-ride native context for such rare cases as COGAGULA.

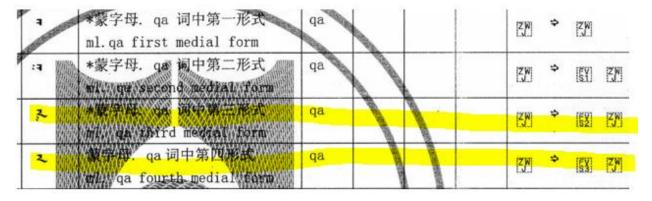


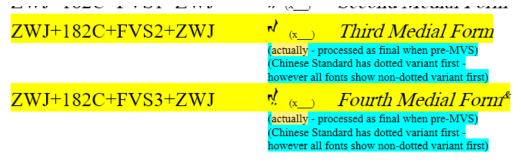
14.) New FVS1 assignment at <U+1843><U+180B> This variant has been attested to.

15.) New FVS3 assignment at <U+1887><U+180F> This is currently listed as an isolate in the Chinese Standard. Professor Quejingzhabu says that this is a mistake and is actually a final.

#### VARIATION BETWEEN CHINESE STANDARD & NEW PROPOSAL

1.) Variation at <U+182C><U+180C> and <U+182C><U+180D>. NP has the undotted form first and the dotted form second. Chinese Standard has them the opposite way. All font developers implement by the NP statement with the undotted form first.





2.) Variation at <U+182D><U+180B> The Chinese Standard uses the non-standard ZWJ to specify an isolate. As this is not necessary and breaks the common paradigm, the NP uses a FVS1.

√ (x\_) Third Medial Form

# FVS MisMatch

| 1.) U+1820-A (post-NNBSP) by the Chir    |                                   |                               |
|--|-----------------------------------|-------------------------------|
| ZWJ+1820+FVS2+ZWJ                        | 1 <sub>(xx)</sub>                 | Third Medial Forn             |
| 1.) U+1820-A (post-NNBSP) by the Univ    | versal Shaping l                  | Engine (Uniscribe) / Harfbuzz |
| 1820+FVS1+ZWJ                            | 1 (xx_)                           | Second Initial Form           |
|  |                                   |                               |
| 2.) U+1828-N (pre-MVS) by the Chinese    | Standard                          |                               |
| ZWJ+1828+FVS2+ZWJ                        | Λ <sup>(x—)</sup>                 | Third Medial Form             |
| 2.) U+1828-N (pre-MVS) by the University | al Shaping Eng                    | ine (Uniscribe) / Harfbuzz    |
| ZWJ+1828+FVS1                            | ر <sub>(×)</sub>                  | Second Final Form             |
| 3.) U+182C-Q (pre-MVS) #1 by the Chin    | nese Standard                     |                               |
| ZWJ+182C+FVS2+ZWJ                        | رير <sub>.</sub> له               | Third Medial Form             |
| 3.) U+182C-Q (pre-MVS) #1 by the Univ    | versal Shaping l                  | Engine (Uniscribe) / Harfbuz  |
| ZWJ+182C                                 | ( <u></u> )                       | First Final Form              |
|  |                                   |                               |
| 4.) U+182C-Q (pre-MVS) #2 by the Chin    | nese Standard                     |                               |
| ZWJ+182C+FVS3+ZWJ                        | ٽ <sub>ر (</sub> ن <sup>—</sup> ) | Fourth Medial For             |
| 4.) U+182C-Q (pre-MVS) #2 by the Univ    | versal Shaping l                  | Engine (Uniscribe) / Harfbuzz |
| ZWJ+182C+FVS1                            | й <sup>(хТ)</sup>                 | Second Final Form             |
|  |                                   |                               |
| 5.) U+182D-G (pre-MVS) by the Chinese    | e Standard                        |                               |

ZWJ+182D+FVS2+ZWJ

| 5.) U+182D-G (pre-MVS) by the Universal Shaping Engine (Uniscribe) / Harfbuzz |   |                           |
|---|---|---------------------------|
| ZWJ+182D+FVS3   | ਔ, <sup>(x</sup> ¯)                     | Fourth Final Form         |
|   |   |                           |
|   |   |                           |
| 6.) U+1835-J (pre-MVS) by the Chinese S                                       | Standard                                |                           |
| ZWJ+1835+FVS1+ZWJ   | ر ~                                     | Second Medial Form        |
| 6.) U+1835-J (pre-MVS) by the Universal                                       | l Shaping Engir                         | ne (Uniscribe) / Harfbuzz |
| ZWJ+1835+FVS1   | ر "                                     | Second Final Form         |
|   |   |                           |
| 7.) U+1836-Y (pre-MVS) by the Chinese   | Standard                                |                           |
| ZWJ+1836+FVS2+ZWJ   | (× <sup>—</sup> )                       | Third Medial Form         |
| 7.) U+1836-Y (pre-MVS) by the Universal                                       | al Shaping Engi                         | ne (Uniscribe) / Harfbuzz |
| ZWJ+1836  | ι (==================================== | Final Form                |
|   |   |                           |