

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

Proposal: Modification of description of variant selectors and variant sequences

By: Enkhdalai B.
enhd_2000@yahoo.com
Mongolia.

Монгол бичгийн Юникод стандартын үсгийн хэлбэр сонголтын хүснэгтийн мэдээллийн хангалтгүй байдал

Санал: Үсгийн хэлбэр сонголт, хэлбэрийн дарааллын хүснэгтийг өөрчлөх

[illegible][illegible]

ಹೊಸದೇವರಾಯ • ೨೭

CONTENTS

1. English part	2
2. Cyrillic Mongolian part	6
3. Classic Mongolian part	11
4. References	17
5. APPENDIXES.	18

Abstract

Existing description of variant selectors and variant sequences in Mongolian Unicode character table (range 1800-180F) in ISO/IEC 10646:2017 and in Unicode 11.0.0 are having a not comprehensive information concerning the variations of characters, and needs some modification.

In this document, there is shown and explained the imperfection of existing description and possible modification of the description of variant selectors and variant sequences is offered.

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.

Note:

The author chose names of variation's category and some other formulations in an arbitrary manner.

1. Basic simple variations: The character has only three forms of variations – initial, medial and final.

For example, 182E ᠮ MONGOLIAN LETTER MA. This letter has only three variations.

182E ᠮ MONGOLIAN LETTER MA

→ 043C м cyrillic small letter em

~ 182E ᠮ᠎ᠠ first form (initial)

~ 182E ᠮ᠎ᠠᠨ first form (medial)

~ 182E ᠮ᠎ᠠᠨᠠ first form (final)

Existing description of variant selectors and variant sequences has enough information.

2. Basic simple variations with additional variations: The character has only three forms of variations – initial, medial and final, that selected automatically by the font. Each of positional variations might be having additional more variations that selected using the format controls FVS1, FVS2 and FVS3.

For example, 1830 ᠰ MONGOLIAN LETTER SA. This letter has only three basic positional variations that selected automatically by the font. The final positional variation of this letter has another two additional variant that selected by FVS1 and FVS2 accordingly.

1830 ᠰ MONGOLIAN LETTER SA

→ 0441 с cyrillic small letter es

- ~ 1830  first form (initial)
- ~ 1830  first form (medial)
- ~ 1830  first form (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: The character has more than one variant that are selected automatically by the font in the positions of initial, medial and final.

For example, 182D  MONGOLIAN LETTER GA. This letter has two initial variations, three medial variations and one final variation, which are selected automatically by the font.

- 182D  MONGOLIAN LETTER GA
- 0433  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 180D  fourth form (medial)

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

Solution of problems and modification

Suggestion on modification of existing description of variant selectors and variant sequences.

- a) In the initial position of the word, the letter has two variations that are selected by the font depend on context. These are  and  , which are shown in following table.

гөлүгэ (göläge)	гараг (γaraγ)
	

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

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

~ 182D ᠭᠣᠯᠦᠭᠦ first form (initial)

~ 182D ᠭᠣᠯᠦᠭᠦ first form (initial)

- b) In the medial position of the word there are three variations of the letter, which will be selected by the font depend on context.

These are ᠲᠠᠷᠠᠫᠤ, ᠲᠠᠷᠠᠭᠠᠷ and ᠲᠠᠷᠠᠭᠠᠷ, which are shown in following table.

таар (taγar)	хагда (qayda)	хэрэп (keger)
		

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

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

~ 182D ᠲᠠᠷᠠᠭᠠᠷ first form (medial)

~ 182D ᠲᠠᠷᠠᠭᠠᠷ first form (medial)

~ 182D ᠲᠠᠷᠠᠭᠠᠷ first form (medial)

- c) Due to the existence of several default forms on initial, medial or final positions, the way of static definition of free variable selectors with the code points is not a good idea.

Therefore, if the letter has only one variation in the position, then shall be used static definition of free variation selectors, as below:

~ 1830 180B ✓ second form (final)

~ 1830 180C ʘ third form (final)

If the letter has more than one variation in the position, then shall be used “**glyph-toggling**” description of free variation selectors, as below:

~ 180B ʘ second form (medial)

~ 180B ʘ second form (medial)

~ 182D 180C ʘ - third form (medial)**

** - third form (medial) can be defined statically to the code point for the letter and not depends on default position (medial) variation forms, i.e. ʘ and ʘ.

d) And finally the modified description of variant sequences can be presented as:

182D ʘ MONGOLIAN LETTER GA

→ 0433 r cyrillic small letter ghe

• Context driven,

~ 182D ʘ first form (initial)

~ 182D ʘ first form (initial)

• Context driven,

~ 182D ʘ first form (medial)

~ 182D ʘ first form (medial)

~ 182D ʘ first form (medial)

• Context driven,

~ 182D ʘ first form (final)

• Positional variations selected by FVS

~ 182D 180B ʘ second form (initial)

~~~ 182D 180B ʘ second form (medial)~~

~ 180B ʘ second form (medial)

~ 180B ʘ second form (medial)

~ 182D 180B ʘ second form (final)

~ 182D 180C ʘ - third form (medial)

~ 182D 180D ʘ fourth form (medial)

####

## Товч

Олон улсын стандарт ISO/IEC 10646:2017 ба Unicode 11.0.0 дах монгол бичгийн кодлолын хүснэгт нь (1800-180F хооронд) монгол бичгийн үсгийн дүрсийн хувирлуудыг бүрэн дүүрэн тусгаж харуулахгүй байгаа тул зарим өөрчлөлт хийх шаардлагатай байна.

Энэхүү баримт бичигт одоо байгаа тайлбарын дутуу талыг үзүүлээд үсгийн хэлбэр сонголт, хэлбэрийн дарааллын хүснэгтийг хэрхэн өөрчилж болох талаар өөрийн саналыг гаргасан болно.

## Үсгийн хэлбэр сонголтын хүснэгтийн мэдээлэл хангалтгүй байдал

Монгол бичгийн үсгийн хэлбэрийн өөрчлөлтийн тайлбарыг олон сонголттой төвөгтэй эсэхийг харгалзан гурван категорид хуваан үзэж болно.

### Жич:

Энд хэрэглэгдэж байгаа категорийн нэр ба зарим бусад томъёоллыг зохиогч өөрөө сонгож нэрлэсэн болно.

1. Энгийн үндсэн хэлбэрүүд: Тухайн үсэг нь үгийн эхэнд, дунд, эцэст гэсэн гуравхан төрлийн хэлбэртэй байна.  
Жишээ нь: 182E ᠮ MONGOLIAN LETTER MA. Монгол бичгийн “м” үсэг нь зөвхөн гурван хэлбэртэй.

182E ᠮ MONGOLIAN LETTER MA

→ 043C м cyrillic small letter em

~ 182E ᠮ first form (initial)

~ 182E ᠮ first form (medial)

~ 182E ᠮ first form (final)

Одоо байгаа үсгийн хэлбэр сонгох хүснэгтийн тайлбарт хангалттай мэдээлэл агуулагдсан байна.

2. Үсгийн нэмэлт хувилбар бүхий энгийн үндсэн хэлбэрүүд: Тухайн үсэг нь үгийн эхэнд, дунд, эцэст гэсэн гурван төрлийн хэлбэртэй ба түүнийг үсгийн тиг автоматаар сонгож харуулна.  
Үсгийн байршил тус бүр нь удирдах тэмдэгт FVS1, FVS2, FVS3 –уудаар өөрчилж болох нэмэлт хувилбаруудтай байна.

Жишээ нь: 1830 ᠰ MONGOLIAN LETTER SA. Монгол бичгийн “с” үсэг нь үсгийн тиг автоматаар сонгодог эхэн, дунд, эцэст орох үндсэн гурван хэлбэртэй.

Энэ үсгийн үгийн эцэст орох хэлбэр нь FVS1 ба FVS2 гэсэн удирдах тэмдэгтийн тусламжтайгаар өөрчилж болох нэмэлт хоёр хэлбэртэй.

1830 𐠰 MONGOLIAN LETTER SA

→ 0441 с cyrillic small letter es

~ 1830 𐠰 first form (initial)

~ 1830 𐠰 first form (medial)

~ 1830 𐠰 first form (final)

~ 1830 180B 𐠰 second form (final)

~ 1830 180C 𐠰 third form (final)

Одоо байгаа үсгийн хэлбэр сонгох хүснэгтийн тайлбарт хангалттай мэдээлэл агуулагдсан байна.

3. Нарийн төвөгтэй хэлбэрүүд: Тухайн үсэг нь үгийн эхэн, дунд, эцсийн байршлуудад үсгийн тиг автоматаар сонгодог хэд хэдэн ялгаатай хэлбэртэй байна.

Жишээ нь: 182D 𐠰 MONGOLIAN LETTER GA. Монгол бичгийн “г” үсэг нь үгийн бүтцээс хамаарч дүрслэгддэг үгийн эхэнд хоёр хэлбэр, үгийн дунд гурван хэлбэртэй байна. Тухайн хэлбэрүүдийг үсгийн тиг автоматаар сонгож гаргана.

182D 𐠰 MONGOLIAN LETTER GA

→ 0433 г cyrillic small letter ghe

~ 182D 𐠰 first form (initial)

~ 182D 𐠰 first form (medial)

~ 182D 𐠰 first form (final)

~ 182D 180B 𐠰 second form (initial)

~ 182D 180B 𐠰 second form (medial)

~ 182D 180B 𐠰 second form (final)

~ 182D 180C 𐠰 third form (medial)

~ 182D 180D 𐠰 fourth form (medial)

Одоо байгаа үсгийн хэлбэр сонгох хүснэгтийн тайлбарын мэдээлэл хангалтгүй, дутуу.

### Асуудлын шийдэл, өөрчлөлт

Үсгийн хэлбэр сонгох, хэлбэрийн дарааллын одоо байгаа тайлбарыг өөрчлөх тухай санал.



- а) Үгийн эхний байрлалд тухайн “г” үсэг нь үсгийн тиг автоматаар сонгож гаргадаг хоёр үндсэн хэлбэртэй байна. Энэ нь:  ба  ба дараах хүснэгтэд харуулав.

|                                                                                   |                                                                                   |
|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| гөлүгэ<br>(gölüge)                                                                | гараг<br>(γaraγ)                                                                  |
|  |  |

Одоо байгаа үсгийн хэлбэр сонгох хүснэгтийн тайлбарт зөвхөн ганц хувилбарыг харуулсан: .

Санал болгох өөрчлөлт нь үгийн эхэнд орох хоёр хувилбарыг хоюуланг харуулах ба дараах хэлбэртэй байна.

~ 182D  first form (initial)

~ 182D  first form (initial)

- б) Үгийн дунд орох “г” үсэг нь үсгийн тиг автоматаар сонгож гаргадаг үндсэн гурван хэлбэртэй байна.

Эдгээр нь , ,  жишээг доорх хүснэгтэд харуулав.

|                                                                                     |                                                                                     |                                                                                     |
|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| тагар<br>(taγar)                                                                    | хагда<br>(qaγda)                                                                    | хэгэр<br>(keger)                                                                    |
|  |  |  |

Одоо байгаа үсгийн хэлбэр сонгох хүснэгтийн тайлбарт зөвхөн ганц хувилбарыг харуулсан: .

Санал болгох өөрчлөлт нь үгийн дунд орох гурван хувилбарыг гурвууланг нь харуулах ба дараах хэлбэртэй байна.

~ 182D  first form (medial)

~ 182D  first form (medial)

~ 182D  first form (medial)

- с) Үгийн эхэн, дунд, эцэст гарах хэд хэдэн хувилбартайн улмаас үсгийн хэлбэр өөрчлөх тэмдэгтүүдийг (FVS1, FVS2, FVS3) үсгийн кодын байрлалтай шууд статик байдлаар оноож өгөх нь тохиромжтой сайн арга биш юм.

Иймээс үгийн доторх байрлалаас хамаарсан үсгийн хувилбар тухайн байрлал тус бүрт ганц байх юм бол хувилбар сонгох тусгай удирдах тэмдэгтийг үсгийн кодын байрлалтай шууд холбон статик байдлаар оноож, дүрсэлж өгч болно. Жишээг доор харуулав:

~ 1830 180B ✓ second form (final)

~ 1830 180C ⇨ third form (final)

Харин тухайн үсэг нь үгийн доторх байрлал тус бүрийн хувьд автоматаар сонгогдож гардаг хэд хэдэн үндсэн хэлбэртэй бол “**glyph-toggling**” буюу “**дүрс-солих**” аргачлалыг үсгийн хувилбар өөрчлөх удирдах тэмдэгийн хэрэглээний тайлбарт ашиглах нь зүйтэй юм. Жишээг доор харуулав:

~ 180B ⇨ second form (medial)

~ 180B ⇨ second form (medial)

~ 182D 180C ⇨ third form (medial)\*\*

\*\* - (үгийн дундах) үсгийн гуравдугаар хувилбарын тухайд тухайн үсгийн хэлбэрээс хамааралгүйгээр (ө.х. 180 ба 180B) удирдах тэмдэгтийг үсгийн кодын байрлалтай статик байдлаар оноож, харуулж болно.

d) Үсгийн хэлбэр сонгох ба хэлбэрийн дарааллын тайлбарын хүснэгтийн өөрчлөлтийг эцсийн байдлаар харуулбал:

182D ⇨ MONGOLIAN LETTER GA

→ 0433 г cyrillic small letter ghe

• Context driven,

~ 182D ⇨ first form (initial)

~ 182D ⇨ first form (initial)

• Context driven,

~ 182D ⇨ first form (medial)

~ 182D ⇨ first form (medial)

~ 182D ⇨ first form (medial)

• Context driven,

~ 182D ⇨ first form (final)

• Positional variations selected by FVS

~ 182D 180B ⇨ second form (initial)

~~~ 182D 180B ⇨ second form (medial)~~

~ 180B ⇨ second form (medial)

~ 180B 𐤁 second form (medial)

~ 182D 180B 𐤂 second form (final)

~ 182D 180C 𐤃 third form (medial)

~ 182D 180D 𐤄 fourth form (medial)

###

сэрлэгдсэн үсэг, гэдсэр, гэдсэрсийгээр
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд

г) гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд

| | |
|--------------------|-----------|
| гөлүгэ
(gölüge) | гэдсэргэд |
| гарар
(γaray) | гэдсэргэд |

гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд

~ 182D first form (initial)
 ~ 182D first form (initial)

г) гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд

| | | |
|------------------|------------------|------------------|
| гарар
(tayar) | хагда
(qayda) | хэгэр
(keger) |
| гэдсэргэд | гэдсэргэд | гэдсэргэд |

гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд
 гэдсэргэд үсэг, гэдсэргэд, гэдсэргэд, гэдсэргэд

~ 182D first form (medial)
 ~ 182D first form (medial)
 ~ 182D first form (medial)

دلیل: ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

~ 1830 180B second form (final)

~ 1830 180C third form (final)

دلیل: ریشه‌یابی در معنی و کاربرد

d

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

~ 180B second form (medial)

~ 180B second form (medial)

~ 182D 180C third form (medial)**

دلیل: ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

• ریشه‌یابی در معنی و کاربرد

182D MONGOLIAN LETTER GA

→ 0433 r cyrillic small letter ghe

• Context driven,

~ 182D first form (initial)

- ~ 182D first form (initial)
- Context driven,
- ~ 182D first form (medial)
- ~ 182D first form (medial)
- ~ 182D first form (medial)
- Context driven,
- ~ 182D first form (final)
- Positional variations selected by FVS
- ~ 182D 180B second form (initial)
- ~~~ 182D 180B second form (medial)~~
- ~ 180B second form (medial)
- ~ 180B second form (medial)
- ~ 182D 180B second form (final)
- ~ 182D 180C third form (medial)
- ~ 182D 180D fourth form (medial)

###

References

1. International Standard ISO/IEC 10646, Fifth Edition, 2017-12, page 207 – 209.
2. Unicode Standard Unicode® 11.0.0, Code Charts, Mongolian range 1800 – 18AF.
3. In the Mongolian writing part of this document there was used a custom developed font “Mng Usug Narrow”, which is created by the author.

Ашиглагдсан хэвлэл, холбоосууд

1. Олон улсын стандарт ISO/IEC 10646, тав дах хувилбар, 2017-12, хуудас 207 – 209.
2. Юникод стандарт Unicode® 11.0.0, кодын хүснэгтүүд, монгол бичгийн хэсэг 1800 – 18AF.
3. Энэхүү баримт бичгийн монгол бичгээр бэлтгэсэн хэсэгт зохиогчийн өөрийн хийсэн тусгай үсгийн тиг “Mng Usug Narrow” –ийг ашигласан болно.

Narrow» Ինչպես նաև «Wide» և «Unicode» ֆոնտերը:

1. Ինչպես օգտագործել «Narrow» ֆոնտը:
2. Ինչպես օգտագործել «Wide» ֆոնտը:
3. Ինչպես օգտագործել «Unicode» ֆոնտը:

###

APPENDIXES

ХАВСРАЛТУУД

שְׁמִי יִשְׂרָאֵל

Modified Description of variant selectors and variant sequences in Mongolian Unicode character table

| No | Code Point | Char. | Existing Description | Modified Description |
|----|------------|-----------|----------------------|----------------------|
| 1 | 1800 | ᠪᠢᠷᠭ᠎ᠠ | MONGOLIAN BIRGA | MONGOLIAN BIRGA |
| 2 | 1801 | ᠡᠭᠢᠯᠢᠯᠢᠰᠢ | MONGOLIAN ELLIPSIS | MONGOLIAN ELLIPSIS |
| | | | | |

(will be added later a complete table)