

DEBORAH
ANDERSON,
SCRIPT ENCODING
INITIATIVE,
UC BERKELEY

ULAANBAATAR,
MONGOLIA
APRIL 3-5, 2019

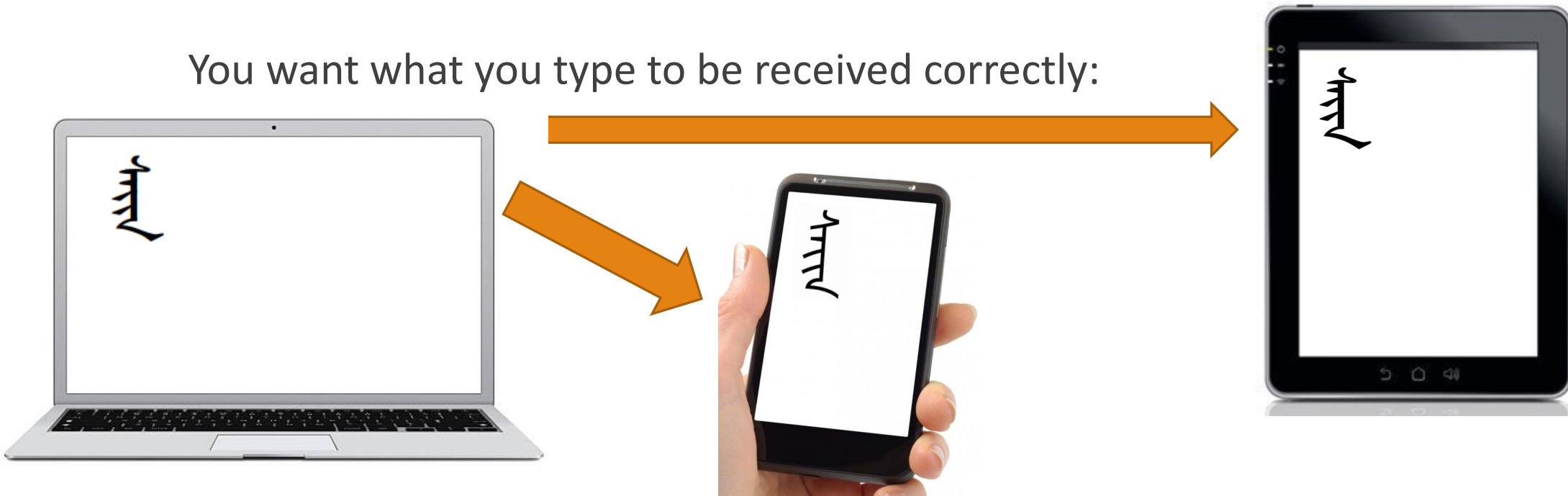
The Unicode Standard and the Unicode Approval Process

Background on the Unicode Standard

Makes it possible so

- text and documents are interchangeable
- platforms supporting Mongolian are interoperable

You want what you type to be received correctly:



Background on the Unicode Standard

What users don't want...

You type:



Users get instead:



Unicode Basics

Unicode assigns a number (“code point”) to the letters and symbols in a script. It also includes a representative glyph and name.



← Glyph

1820

← Code point

MONGOLIAN LETTER A ← Name



Unicode Basics

Code chart

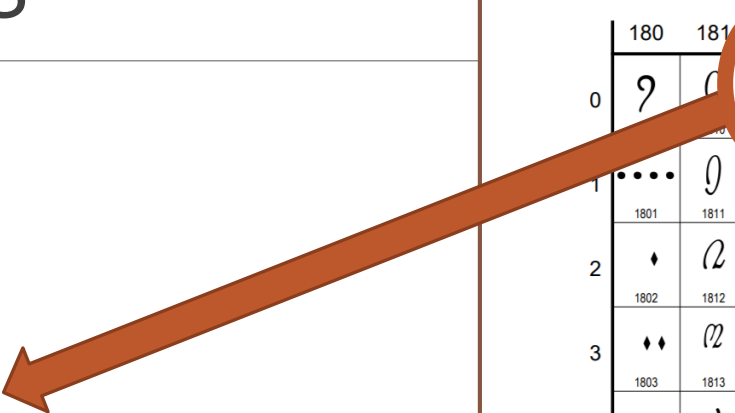
1800		Mongolian										18AF
		180	181	182	183	184	185	186	187	188	189	18A
0		ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠᠨ
1		ᠠᠨ	ᠡᠨ	ᠢᠨ	ᠣᠨ	ᠤᠨ	ᠥᠨ	ᠦᠨ	ᠨᠠ	ᠨᠡ	ᠨᠢ	ᠨᠣ
2		ᠠᠨᠠ	ᠡᠨᠠ	ᠢᠨᠠ	ᠣᠨᠠ	ᠤᠨᠠ	ᠥᠨᠠ	ᠦᠨᠠ	ᠨᠠᠨ	ᠨᠡᠨ	ᠨᠢᠨ	ᠨᠣᠨ
3		ᠠᠨᠡ	ᠡᠨᠡ	ᠢᠨᠡ	ᠣᠨᠡ	ᠤᠨᠡ	ᠥᠨᠡ	ᠦᠨᠡ	ᠨᠠᠨᠠ	ᠨᠡᠨᠠ	ᠨᠢᠨᠠ	ᠨᠣᠨᠠ
4		ᠠᠨᠢ	ᠡᠨᠢ	ᠢᠨᠢ	ᠣᠨᠢ	ᠤᠨᠢ	ᠥᠨᠢ	ᠦᠨᠢ	ᠨᠠᠨᠡ	ᠨᠡᠨᠡ	ᠨᠢᠨᠡ	ᠨᠣᠨᠡ
5		ᠠᠨᠣ	ᠡᠨᠣ	ᠢᠨᠣ	ᠣᠨᠣ	ᠤᠨᠣ	ᠥᠨᠣ	ᠦᠨᠣ	ᠨᠠᠨᠢ	ᠨᠡᠨᠢ	ᠨᠢᠨᠢ	ᠨᠣᠨᠢ
6		ᠠᠨᠤ	ᠡᠨᠤ	ᠢᠨᠤ	ᠣᠨᠤ	ᠤᠨᠤ	ᠥᠨᠤ	ᠦᠨᠤ	ᠨᠠᠨᠣ	ᠨᠡᠨᠣ	ᠨᠢᠨᠣ	ᠨᠣᠨᠣ
7		ᠠᠨᠥ	ᠡᠨᠥ	ᠢᠨᠥ	ᠣᠨᠥ	ᠤᠨᠥ	ᠥᠨᠥ	ᠦᠨᠥ	ᠨᠠᠨᠤ	ᠨᠡᠨᠤ	ᠨᠢᠨᠤ	ᠨᠣᠨᠤ
8		ᠠᠨᠦ	ᠡᠨᠦ	ᠢᠨᠦ	ᠣᠨᠦ	ᠤᠨᠦ	ᠥᠨᠦ	ᠦᠨᠦ	ᠨᠠᠨᠥ	ᠨᠡᠨᠥ	ᠨᠢᠨᠥ	ᠨᠣᠨᠥ
9		ᠠᠨᠨ	ᠡᠨᠨ	ᠢᠨᠨ	ᠣᠨᠨ	ᠤᠨᠨ	ᠥᠨᠨ	ᠦᠨᠨ	ᠨᠠᠨᠦ	ᠨᠡᠨᠦ	ᠨᠢᠨᠦ	ᠨᠣᠨᠦ
A		ᠠᠨᠨᠠ	ᠡᠨᠨᠠ	ᠢᠨᠨᠠ	ᠣᠨᠨᠠ	ᠤᠨᠨᠠ	ᠥᠨᠨᠠ	ᠦᠨᠨᠠ	ᠨᠠᠨᠨ	ᠨᠡᠨᠨ	ᠨᠢᠨᠨ	ᠨᠣᠨᠨ

<https://www.unicode.org/charts/PDF/U1800.pdf>



Unicode Basics

Code chart



1800		Mongolian										18AF
	180	181	182	183	184	185	186	187	188	189	18A	
0	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠᠨ	
1	ᠠᠨ	ᠡᠨ	ᠢᠨ	ᠣᠨ	ᠤᠨ	ᠥᠨ	ᠦᠨ	ᠨᠠ	ᠨᠡ	ᠨᠢ	ᠨᠣ	
2	ᠠᠨᠠ	ᠠᠨᠡ	ᠠᠨᠢ	ᠠᠨᠣ	ᠠᠨᠤ	ᠠᠨᠥ	ᠠᠨᠦ	ᠠᠨᠨ	ᠠᠨᠬ	ᠠᠨᠭ	ᠠᠨᠠᠨ	
3	ᠠᠨᠠᠨ	ᠠᠨᠠᠨᠠ	ᠠᠨᠠᠨᠡ	ᠠᠨᠠᠨᠢ	ᠠᠨᠠᠨᠣ	ᠠᠨᠠᠨᠤ	ᠠᠨᠠᠨᠥ	ᠠᠨᠠᠨᠦ	ᠠᠨᠠᠨᠨ	ᠠᠨᠠᠨᠬ	ᠠᠨᠠᠨᠭ	
4	ᠠᠨᠠᠨᠠ	ᠠᠨᠠᠨᠠᠠ	ᠠᠨᠠᠨᠠᠡ	ᠠᠨᠠᠨᠠᠢ	ᠠᠨᠠᠨᠠᠣ	ᠠᠨᠠᠨᠠᠤ	ᠠᠨᠠᠨᠠᠥ	ᠠᠨᠠᠨᠠᠦ	ᠠᠨᠠᠨᠠᠨ	ᠠᠨᠠᠨᠠᠬ	ᠠᠨᠠᠨᠠᠭ	
5	ᠠᠨᠠᠨᠠᠨ	ᠠᠨᠠᠨᠠᠨᠠ	ᠠᠨᠠᠨᠠᠨᠡ	ᠠᠨᠠᠨᠠᠨᠢ	ᠠᠨᠠᠨᠠᠨᠣ	ᠠᠨᠠᠨᠠᠨᠤ	ᠠᠨᠠᠨᠠᠨᠥ	ᠠᠨᠠᠨᠠᠨᠦ	ᠠᠨᠠᠨᠠᠨᠨ	ᠠᠨᠠᠨᠠᠨᠬ	ᠠᠨᠠᠨᠠᠨᠭ	
6	ᠠᠨᠠᠨᠠᠨᠠ	ᠠᠨᠠᠨᠠᠨᠠᠠ	ᠠᠨᠠᠨᠠᠨᠠᠡ	ᠠᠨᠠᠨᠠᠨᠠᠢ	ᠠᠨᠠᠨᠠᠨᠠᠣ	ᠠᠨᠠᠨᠠᠨᠠᠤ	ᠠᠨᠠᠨᠠᠨᠠᠥ	ᠠᠨᠠᠨᠠᠨᠠᠦ	ᠠᠨᠠᠨᠠᠨᠠᠨ	ᠠᠨᠠᠨᠠᠨᠠᠬ	ᠠᠨᠠᠨᠠᠨᠠᠭ	
7	ᠠᠨᠠᠨᠠᠨᠠᠨ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠ	ᠠᠨᠠᠨᠠᠨᠠᠨᠡ	ᠠᠨᠠᠨᠠᠨᠠᠨᠢ	ᠠᠨᠠᠨᠠᠨᠠᠨᠣ	ᠠᠨᠠᠨᠠᠨᠠᠨᠤ	ᠠᠨᠠᠨᠠᠨᠠᠨᠥ	ᠠᠨᠠᠨᠠᠨᠠᠨᠦ	ᠠᠨᠠᠨᠠᠨᠠᠨᠨ	ᠠᠨᠠᠨᠠᠨᠠᠨᠬ	ᠠᠨᠠᠨᠠᠨᠠᠨᠭ	
8	ᠠᠨᠠᠨᠠᠨᠠᠨᠠ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠠ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠡ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠢ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠣ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠤ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠥ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠦ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠬ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠭ	
9	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠠ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠡ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠢ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠣ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠤ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠥ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠦ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠨ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠬ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠭ	
A	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠠ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠠᠠ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠠᠡ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠠᠢ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠠᠣ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠠᠤ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠠᠥ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠠᠦ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠠᠬ	ᠠᠨᠠᠨᠠᠨᠠᠨᠠᠨᠠᠭ	

<https://www.unicode.org/charts/PDF/U1800.pdf>

Unicode Basics

Names list

1800

Mongolian

1825

Punctuation

Additional birgas are encoded in the Mongolian Supplement block at 11660-1167F.

- 1800 ʘ MONGOLIAN BIRGA
→ 0F04 tibetan mark initial yig mgo mdun ma
- 1801 ⋯ MONGOLIAN ELLIPSIS
- 1802 ∙ MONGOLIAN COMMA
- 1803 ∙∙ MONGOLIAN FULL STOP
- 1804 ∶ MONGOLIAN COLON
- 1805 ⋄ MONGOLIAN FOUR DOTS
• marks the end of a chapter
• not a format control character, but simply a hyphen for Todo
• also used in the Hudum (modern Mongolian) writing system
→ 00AD soft hyphen
- 1807 † MONGOLIAN SIBE SYLLABLE BOUNDARY MARKER
- 1808 › MONGOLIAN MANCHU COMMA
- 1809 ›› MONGOLIAN MANCHU FULL STOP
- 180A - MONGOLIAN NIRUGU

Format controls

- 180B 𐎆 MONGOLIAN FREE VARIATION SELECTOR ONE
• abbreviated FVS1
- 180C 𐎇 MONGOLIAN FREE VARIATION SELECTOR TWO
• abbreviated FVS2
- 180D 𐎈 MONGOLIAN FREE VARIATION SELECTOR THREE
• abbreviated FVS3
- 180E 𐎉 MONGOLIAN VOWEL SEPARATOR
• abbreviated MVS

Digits

- 1810 0 MONGOLIAN DIGIT ZERO
- 1811 1 MONGOLIAN DIGIT ONE
- 1812 2 MONGOLIAN DIGIT TWO
- 1813 3 MONGOLIAN DIGIT THREE
- 1814 4 MONGOLIAN DIGIT FOUR
- 1815 5 MONGOLIAN DIGIT FIVE
- 1816 6 MONGOLIAN DIGIT SIX
- 1817 7 MONGOLIAN DIGIT SEVEN
- 1818 8 MONGOLIAN DIGIT EIGHT
- 1819 9 MONGOLIAN DIGIT NINE

Basic letters


Cross-references show commonly used Cyrillic transliterations.







Letters unspecified for distinct positional forms are shown without them. Letters with unspecified positional forms use a squared question mark to indicate the gap. All vowels are expected to have explicit isolate forms. For consonants the isolate first form is the initial first form unless specified otherwise.

- 1820 ᠠ MONGOLIAN LETTER A
→ 0430 a cyrillic small letter a
~ 1820 ᠠ᠊ first form (isolate)
~ 1820 ᠡ᠊ first form (initial)
~ 1820 ᠢ᠊ first form (medial)
~ 1820 ᠣ᠊ first form (final)
~ 1820 180B ᠡᠭ second form (isolate)
~ 1820 180B ᠢᠭ second form (medial)
~ 1820 180B ᠣᠭ second form (final)
~ 1820 180C ᠠᠨ third form (medial)
- 1821 ᠡ MONGOLIAN LETTER E
→ 044D э cyrillic small letter e
~ 1821 ᠡ᠊ first form (isolate)
~ 1821 ᠢ᠊ first form (initial)
~ 1821 ᠣ᠊ first form (medial)
~ 1821 ᠣ᠊ first form (final)
~ 1821 180B ᠡᠭ second form (initial)
~ 1821 180B ᠣᠭ second form (final)
- 1822 ᠢ MONGOLIAN LETTER I
→ 043B и cyrillic small letter i
~ 1822 ᠢ᠊ first form (isolate)
~ 1822 ᠣ᠊ first form (initial)
~ 1822 ᠣ᠊ first form (medial)
~ 1822 ᠣ᠊ first form (final)
~ 1822 180B ᠢᠭ second form (medial)
- 1823 ᠣ MONGOLIAN LETTER O
→ 043E о cyrillic small letter o
~ 1823 ᠣ᠊ first form (isolate)
~ 1823 ᠣ᠊ first form (initial)
~ 1823 ᠣ᠊ first form (medial)
~ 1823 ᠣ᠊ first form (final)
~ 1823 180B ᠣᠭ second form (medial)
~ 1823 180B ᠣᠭ second form (final)
- 1824 ᠤ MONGOLIAN LETTER U
→ 0443 у cyrillic small letter u
~ 1824 ᠤ᠊ first form (isolate)
~ 1824 ᠣ᠊ first form (initial)
~ 1824 ᠣ᠊ first form (medial)
~ 1824 ᠣ᠊ first form (final)
~ 1824 180B ᠣᠭ second form (medial)
- 1825 ᠥ MONGOLIAN LETTER OE
→ 04E9 ө cyrillic small letter barred o
~ 1825 ᠥ᠊ first form (isolate)
~ 1825 ᠣ᠊ first form (initial)
~ 1825 ᠣ᠊ first form (medial)
~ 1825 ᠣ᠊ first form (final)
~ 1825 180B ᠥᠭ second form (medial)
~ 1825 180B ᠥᠭ second form (final)
~ 1825 180C ᠥᠨ third form (medial)

Unicode Basics

Names list

1801	----	MONGOLIAN ELLIPSIS
1802	⋅	MONGOLIAN COMMA
1803	⋅⋅	MONGOLIAN FULL STOP
1804	:	MONGOLIAN COLON
1805	⋄	MONGOLIAN FOUR DOTS
		• marks the end of a chapter
1806	-	MONGOLIAN TODO SOFT HYPHEN
		• not a format control character, but simply a hyphen for Todo
		• also used in the Hudum (modern Mongolian) writing system
		→ 00AD  soft hyphen
1807	⤵	MONGOLIAN SIBE SYLLABLE BOUNDARY MARKER
1808	›	MONGOLIAN MANCHU COMMA
1809	››	MONGOLIAN MANCHU FULL STOP
180A	-	MONGOLIAN NIRUGU

1800	Mongolian	1825
Punctuation		
Additional birgas are encoded in the Mongolian Supplement block at 11660-1167F.		
1800	ᠰ	MONGOLIAN BIRGA → 0F04  tibetan mark initial yig mgo mdun
1801	⋯	MONGOLIAN ELLIPSIS
1802	⋅	MONGOLIAN COMMA
1803	⋅⋅	MONGOLIAN FULL STOP
1804	:	MONGOLIAN COLON
1805	⋄	MONGOLIAN FOUR DOTS
		• marks the end of a chapter
1806	-	MONGOLIAN TODO SOFT HYPHEN
		• not a format control character, but simply a hyphen for Todo
		• also used in the Hudum (modern Mongolian) writing system
		→ 00AD  soft hyphen
1807	⤵	MONGOLIAN SIBE SYLLABLE BOUNDARY MARKER
1808	›	MONGOLIAN MANCHU COMMA
1809	››	MONGOLIAN MANCHU FULL STOP
180A	-	MONGOLIAN NIRUGU
Format controls		
180B		MONGOLIAN FREE VARIATION SELECTOR ONE • abbreviated FVS1
180C		MONGOLIAN FREE VARIATION SELECTOR TWO • abbreviated FVS2
180D		MONGOLIAN FREE VARIATION SELECTOR THREE • abbreviated FVS3
180E		MONGOLIAN VOWEL SEPARATOR • abbreviated MVS
Digits		
1810	᠐	MONGOLIAN DIGIT ZERO
1811	᠑	MONGOLIAN DIGIT ONE
1812	᠒	MONGOLIAN DIGIT TWO
1813	᠓	MONGOLIAN DIGIT THREE
1814	᠔	MONGOLIAN DIGIT FOUR
1815	᠕	MONGOLIAN DIGIT FIVE
1816	᠖	MONGOLIAN DIGIT SIX
1817	᠗	MONGOLIAN DIGIT SEVEN
1818	᠘	MONGOLIAN DIGIT EIGHT
1819	᠙	MONGOLIAN DIGIT NINE
Basic letters		
Cross-references show commonly used Cyrillic transliterations.		
Letters unspecified for distinct positional forms are shown without them. Letters with unspecified positional forms use a squared question mark to indicate the gap. All vowels are expected to have explicit isolate forms. For consonants the isolate first form is the initial first form unless specified otherwise.		
1820	ᠠ	MONGOLIAN LETTER A → 0430 а cyrillic small letter a ~ 1820 ᠠ᠊ first form (isolate) ~ 1820 ᠠ᠊᠊ first form (initial) ~ 1820 ᠠ᠊᠎ first form (medial) ~ 1820 ᠠ᠊᠎᠎ first form (final) ~ 1820 180B ᠠ᠊᠎᠎᠎ second form (isolate) ~ 1820 180B᠎᠎ second form (medial) ~ 1820 180B᠎᠎᠎ second form (final) ~ 1820 180C ᠠ᠊᠎᠎᠎᠎ third form (medial)
1821	ᠡ	MONGOLIAN LETTER E → 044D э cyrillic small letter e ~ 1821 ᠡ᠊ first form (isolate) ~ 1821 ᠡ᠊᠊ first form (initial) ~ 1821 ᠡ᠊᠎ first form (medial) ~ 1821 ᠡ᠊᠎᠎ first form (final) ~ 1821 180B ᠡ᠊᠎᠎᠎ second form (initial) ~ 1821 180B᠎᠎ second form (final)
1822	ᠢ	MONGOLIAN LETTER I → 043B и cyrillic small letter i ~ 1822 ᠢ᠊ first form (isolate) ~ 1822 ᠢ᠊᠊ first form (initial) ~ 1822 ᠢ᠊᠎ first form (medial) ~ 1822 ᠢ᠊᠎᠎ first form (final) ~ 1822 180B ᠢ᠊᠎᠎᠎ second form (medial)
1823	ᠣ	MONGOLIAN LETTER O → 043E о cyrillic small letter o ~ 1823 ᠣ᠊ first form (isolate) ~ 1823 ᠣ᠊᠊ first form (initial) ~ 1823 ᠣ᠊᠎ first form (medial) ~ 1823 ᠣ᠊᠎᠎ first form (final) ~ 1823 180B ᠣ᠊᠎᠎᠎ second form (medial) ~ 1823 180B᠎᠎ second form (final)
1824	ᠤ	MONGOLIAN LETTER U → 0443 у cyrillic small letter u ~ 1824 ᠤ᠊ first form (isolate) ~ 1824 ᠤ᠊᠊ first form (initial) ~ 1824 ᠤ᠊᠎ first form (medial) ~ 1824 ᠤ᠊᠎᠎ first form (final) ~ 1824 180B ᠤ᠊᠎᠎᠎ second form (medial) ~ 1824 180B᠎᠎ second form (final)
1825	ᠥ	MONGOLIAN LETTER OE → 04E9 ө cyrillic small letter barred o ~ 1825 ᠥ᠊ first form (isolate) ~ 1825 ᠥ᠊᠊ first form (initial) ~ 1825 ᠥ᠊᠎ first form (medial) ~ 1825 ᠥ᠊᠎᠎ first form (final) ~ 1825 180B ᠥ᠊᠎᠎᠎ second form (medial) ~ 1825 180B᠎᠎ second form (final) ~ 1825 180C ᠥ᠊᠎᠎᠎᠎ third form (medial)

Unicode Basics

Names list

1820 ᠠ MONGOLIAN LETTER A
→ 0430 а cyrillic small letter a
~ 1820 ᠠ first form (isolate)
~ 1820 ᠠ first form (initial)
~ 1820 ᠠ first form (medial)
~ 1820 ᠠ first form (final)
~ 1820 180B ᠠ second form (isolate)
~ 1820 180B ᠠ second form (medial)
~ 1820 180B ᠠ second form (final)
~ 1820 180C ᠠ third form (medial)

Including variant information in names list poses problems:

- Not the usual approach
- Can't capture all the details of writing system
- Hard to maintain
- Recommend a *different* approach

Unicode Basics

Names list

1820	ᠠ	MONGOLIAN LETTER A
		→ 0430 а cyrillic small letter a
		~ 1820 ᠠ first form (isolate)
		~ 1820 ᠠ second form (medial)
		~ 1820 ᠠ third form (final)
		~ 1820 ᠠ fourth form (isolate)
		~ 1820 ᠠ fifth form (medial)
		~ 1820 ᠠ sixth form (final)
		~ 1820 180C ᠠ third form (medial)

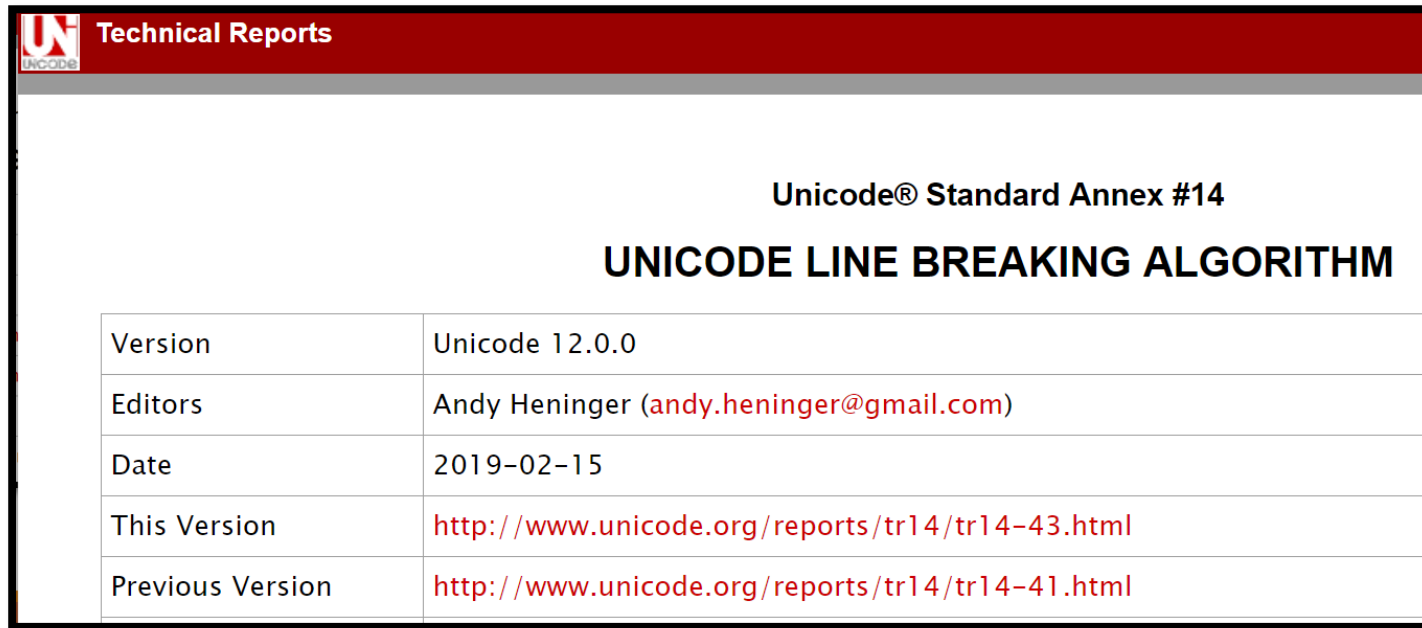
Unicode® Technical Notes

- Separate from Unicode publication and its publishing cycle
- Easy to edit
- Contents could become a Unicode Technical Standard, after community agreement

<https://unicode.org/notes/>

Unicode Basics

The Unicode Standard *also* includes information on characters' behavioral properties (such as line breaking, text segmentation, etc.)

A screenshot of a web page titled "Technical Reports" with the Unicode logo. The main heading is "Unicode® Standard Annex #14" followed by "UNICODE LINE BREAKING ALGORITHM". Below this is a table with five rows: Version (Unicode 12.0.0), Editors (Andy Heninger), Date (2019-02-15), This Version (a URL), and Previous Version (a URL).

Unicode® Standard Annex #14	
UNICODE LINE BREAKING ALGORITHM	
Version	Unicode 12.0.0
Editors	Andy Heninger (andy.heninger@gmail.com)
Date	2019-02-15
This Version	http://www.unicode.org/reports/tr14/tr14-43.html
Previous Version	http://www.unicode.org/reports/tr14/tr14-41.html

<https://www.unicode.org/reports/tr14/>

Character Approval and Support Process

Character Approval Process

Preparation of proposal

Example from a 2017 proposal to add one character:

ᠵ MONGOLIAN LETTER CHA WITH TWO DOTS

Note that proposals should include evidence:

Fig. 2: Tsydendambaev 1972 p. 304

диакритический знак. Правда, стремление как-то разграничить буквы *g* и *k*, омонимичные по своему изображению в классическом монгольском языке (ᠭ · ᠬ · ᠭ), имеет известное основание: в живой речи они приобрели фонематические различия. В "Истории перекочевки в Баргузин..." путем постановки точек справа от знака *ṣādhe* (ᠵ) показано, что монгольская буква *ṣ* в бурятском языке произносится как звук *ṣ̣*.

Character Approval Process

Preparation of proposal



Review by Unicode Script Ad Hoc



- Reviews proposals and makes recommendations to standards committees



Character Approval Process

Preparation of proposal



Review by Unicode Script Ad Hoc



Unicode Technical Committee

- Reviews proposals and makes decisions on proposals

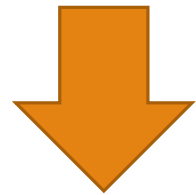


Character Approval Process

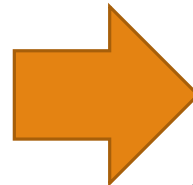
Preparation of proposal



Review by Script Ad Hoc



Unicode Technical Committee



ISO SC2/WG2

- Reviews proposals and decides whether to put on ballot

Character Approval Process

Preparation of proposal



Review by Script Ad Hoc



Unicode Technical Committee



ISO SC2/WG2



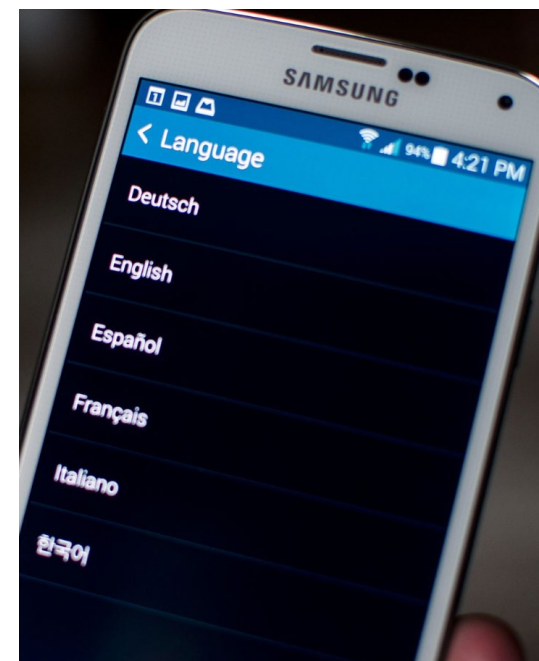
Character Support Process

- After publication need:

Fonts and keyboard definition



Locale data



Character Approval and Support Process - Summary

- Many steps are involved in getting characters approved
- Can take several years for characters to be approved and implemented on devices



- Working together helps the process go smoothly

Character Approval and Support Process - Summary

- Many steps involved in getting characters approved
- Can take several years for characters to be approved and implemented on devices
- Working together helps the process go smoothly

Баярлалаа!