



Script Encoding Working Group

Recommendations to UTC #189 (October 2026) on Script Proposals

Date: 2026-10-23

From: Jan Kučera, Roozbeh Pournader, and Deborah Anderson

Based on notes by: Jan Kučera, Quinn Daedal, Deborah Anderson, and Manish Goregaokar

Also participating: Glenda Bellarosa, Peter Constable, Eiso Chan, Quinn Daedal, Lorna Evans, Andrew Glass, Manish Goregaokar, Liang Hai, Ned Holbrook, Anushah Hossain, John Hudson, Denis Moyogo Jacquerye, Kushim Jiang, Helena Kansa, Robin Leroy, Kirk Miller, Ben Mitchell, Anshuman Pandey, Katie Russell, CheonHyeong Sim, Ivo Spira, Michel Suignard, Harald Tveiten, Ken Whistler, Lawrence Wolf-Sonkin, Ben Yang, and Oreen Yousuf

The Script Encoding Working Group met on August 7, September 4, October 2, and October 9, 2026, in order to review proposals. This document represents feedback on proposals that were available when the group met.

Table of Contents

Topics in bold contain recommendations to the UTC.

1 New Scripts

- 1.1 [Maya Hieroglyphs](#)
- 1.2 [N'ti](#)
- 1.3 [Ndiko Jonam](#)
- 1.4 [Koré Sébèli](#)
- 1.5 [Mwangwego script](#)
- 1.6 [Zou](#)
- 1.7 [Kadu](#)
- 1.8 [Wei Rma](#)
- 1.9 [Wagindara](#)
- 1.10 [Quikscript](#)

2 New Characters

- 2.1 [Arabic Letter Alef Maksura With Comma Above](#)
- 2.2 [Devanagari Sign Final Anusvara](#)
- 2.3 [Three new characters of Devanagari](#)
- 2.4 [Devanagari Correction Mark For Vowel Signs](#)
- 2.5 [Dotless I/i with descender](#)
- 2.6 [Latin Capital Turned W](#)
- 2.7 [Latin Small Letter Turned Dotless I](#)
- 2.8 [Ellis's Phonotypic Alphabet for Yahgan and Ona](#)
- 2.9 [South Slavic Cyrillic modifier letters](#)
- 2.10 [Myanmar letters for Kham Dom Pa'O](#)

- 2.11 [Malayalam Letter Vedic Double Anusvara](#)
- 2.12 [Turkestan Brāhmī](#)
- 2.13 [Arabic Characters from Indo-Pak Mushafs](#)
- 2.14 [Benjamin Franklin's reformed alphabet](#)

3 New Symbols

- 3.1 [Leibniz: 6 letterlike symbols](#)
- 3.2 [Leibniz: Miscellaneous symbols](#)
- 3.3 [Turned Digit One for Balanced Ternary](#)
- 3.4 [Big5 ETen Extension Box Drawing Characters](#)

4 Changes to Characters

- 4.1 [Report on UTC 158-A106](#)
- 4.2 [Report on UTC 164-A63](#)
- 4.3 [Character properties of U+0CD5 KANNADA LENGTH MARK](#)

5 Other

- 5.1 [Vithkuqi annotations](#)
- 5.2 [Report on UTC 159-A81](#)
- 5.3 [On U+055F ARMENIAN ABBREVIATION MARK](#)
- 5.4 [Repha in Kannada](#)
- 5.5 [Nukta in Sinhala](#)
- 5.6 [Cross-Writing-System Ambiguity in Mongolian Script](#)

6 In Process

1. New Scripts

1.1 Maya Hieroglyphs #866

Documents:

[L2/26-145R3](#): Proposal to Encode Codical Maya Hieroglyphs — Dr. Des. Carlos Pallan, et al

Recommendation:

1. **[SEW-189-C1] Consensus** : Provisionally assign 513 code points: U+15500..U+156DC and U+156EC..U+156EF in a new Mayan Hieroglyphs block, and U+156F0..U+1570F in a new Mayan Hieroglyph Format Controls block, as described in L2/26-145R3. [Ref: 1.1 in L2/26-239]
2. **[SEW-189-A1] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 513 code points U+15500..U+156DC, U+156EC..U+1570F, as described in L2/26-145R3. [Ref: 1.1 in L2/26-239]

Comments:

This is a proposal to encode Maya hieroglyphs in Unicode. This document focuses on the postclassic codical subset from 1100–1519 CE, comprising 477 signs, 4 characters for representing empty, lost, erased, and blank signs, and 32 format controls. Akin to the Egyptian hieroglyph format controls, the Mayan hieroglyph format controls allow the signs to be positioned relative to one another, joined, rotated and mirrored, grouped into segments, compressed or marked as damaged. One control character also indicates that the previous sign should be rendered in color.

The corpus of signs was prepared with the scholarly community after extensive research and covers signs in codices of Dresden, Madrid and Paris. The proposal has undergone several iterations in the SEW and is ready to be recommended. Work on a separate proposal extending the repertoire to cover the preceding classic period is ongoing.



1.2 N'ti #768

Documents:

[L2/26-087](#): Proposal for Encoding the N'ti Script in the UCS (supercedes L2/26-028R) — Oreen Yousuf, Ibrahima Ceesay

Related:

[L2/26-158](#): Recommendations to UTC #188 (July 2026) on Script Proposals — SEW / Jan Kučera, et al

Recommendation:

1. **[SEW-189-C2] Consensus** : Provisionally assign 47 code points U+1CB80..U+1CBA4, U+1CBA6..U+1CBAF, in a new Nti block at U+1CB80..U+1CBAF, as described in Section V of L2/26-087. [Ref: 1.2 in L2/26-239]
2. **[SEW-189-A2] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 47 provisionally assigned code points U+1CB80..U+1CBA4, U+1CBA6..U+1CBAF, in a new Nti block: U+1CB80..U+1CBAF, as described in Section V of L2/26-087. [Ref: 1.2 in L2/26-239]
3. **[SEW-189-A3] Action Item** for Debbie Anderson, SAH: Work with the proposal authors to submit a compliant font for N'ti to Unicode. [Ref: 1.2 in L2/26-239]

Comments:

This is a proposal to encode the N'ti script invented in 1966 used to write the Soninke languages of West Africa. This script has been recommended for provisional assignment in L2/26-158, but UTC remanded it back to SEW to ensure legal clearance from the script authors. The SEW now confirms to the best of its knowledge that the script meets Unicode's legal requirements for encoding.

ᠠᠨ ᠠᠨᠠᠨ ᠠᠨᠠᠨ ᠠᠨᠠᠨ
ᠠᠨ ᠠᠨᠠᠨ ᠠᠨᠠᠨ ᠠᠨᠠᠨ ᠠᠨᠠᠨ
(ᠠᠨ) ᠠᠨᠠᠨ ᠠᠨᠠᠨ

1.3 Ndiko Jonam #769

Documents:

[L2/26-055R](#): Proposal to Encode the Ndiko Jonam Script in the UCS — Oreen Yousuf

Related:

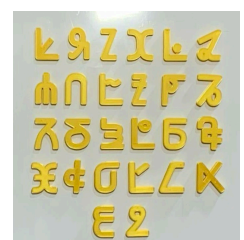
[L2/26-158](#): Recommendations to UTC #188 (July 2026) on Script Proposals — SEW / Jan Kučera, et al

Recommendation:

1. **[SEW-189-C3] Consensus** : Provisionally assign 69 code points U+1CB30..U+1CB49, U+1CB51..U+1CB71, and U+1CB76..U+1CB7F, in a new Ndiko Jonam block at U+1CB30..U+1CB7F, as described in L2/26-055R. [Ref: 1.3 in L2/26-239]
2. **[SEW-189-A4] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 69 provisionally assigned code points U+1CB30..U+1CB49, U+1CB51..U+1CB71, and U+1CB76..U+1CB7F, in a new Ndiko Jonam block: U+1CB30..U+1CB7F, as described in L2/26-055R. [Ref: 1.3 in L2/26-239]
3. **[SEW-189-A5] Action Item** for Debbie Anderson, SAH: Work with the proposal authors to submit a compliant font for Ndiko Jonam to Unicode. [Ref: 1.3 in L2/26-239]

Comments:

This is a proposal for an alphabetical script invented between 2009 and 2012 and used to write the Luo languages of East Africa. This script has been recommended for provisional assignment in L2/26-158, but UTC remanded it back to SEW to ensure legal clearance from the script authors. The SEW now confirms to the best of its knowledge that the script meets Unicode's legal requirements for encoding.



1.4 Koré Sébèli #29

Documents:

[L2/26-132R](#): Proposal for the encoding of « KORE SEBELI » — Mohamed Bentoura Bangoura

Related:

[L2/26-158](#): Recommendations to UTC #188 (July 2026) on Script Proposals — SEW / Jan Kučera, et al

Recommendation:

1. **[SEW-189-C4] Consensus** : Provisionally assign 93 code points U+15B00..U+15B52, U+15B56..U+15B5F, in a new Kore Sebeli block at U+15B00..U+15B5F, described in L2/26-132R, as amended: U+15B1B KORE SEBELI CAPITAL LETTER OU WITH THREE DOTS ABOVE. [Ref: 1.4 in L2/26-239]
2. **[SEW-189-A6] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 93 provisionally assigned code points U+15B00..U+15B52, U+15B56..U+15B5F, in a new Kore Sebeli block: U+15B00..U+15B5F, described in L2/26-132R, as amended. [Ref: 1.4 in L2/26-239]
3. **[SEW-189-A7] Action Item** for Debbie Anderson, SAH: Work with the proposal authors to submit a compliant font for Koré Sébèli to Unicode. [Ref: 1.4 in L2/26-239]

Comments:

This is a proposal for Koré Sébèli, a recently invented script for the Soso language of West Africa. This script has been recommended for provisional assignment in L2/26-158, but UTC remanded it back to SEW to ensure legal clearance from the script authors and revise character names. The SEW now confirms to the best of its knowledge that the script meets Unicode's legal requirements for encoding. The proposal was updated with character names as suggested by SEW (SMALL LETTER instead of LOWERCASE, PLUS SIGN and MINUS SIGN).

Ɔǎhǎǎ Ɔǎǎǎǎ Ɔǎǎǎǎ, Ɔǎǎ Ɔǎǎ Ɔǎǎǎ Ɔǎǎǎǎ
Ɔǎǎǎǎ Ɔǎǎǎǎ Ɔǎǎǎǎ, Ɔǎǎǎǎ Ɔǎǎǎǎ Ɔǎǎǎǎ

1.5 Mwangwego script #543

Related:

[L2/26-010](#): Recommendations to UTC #186 (January 2026) on Script Proposals — SEW / Jan Kučera, et al

[L2/26-029](#): Final Proposal for Encoding the Mwangwego Script in the UCS — Oreen Yousuf, Daniel Yacob

Recommendation:

No action is requested of the UTC.

Comments:

The SEW also confirms that to the best of its knowledge, the Mwangwego script provisionally assigned by 186-C19 now meets Unicode's legal requirements for encoding.

1.6 Zou #910

Documents:

[L2/26-258](#): Proposal to Encode the Zou Script and Numerals in Unicode — Dr. T Marvin Zou

Related:

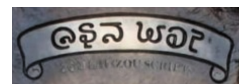
[L2/10-254](#): Introducing the Zou Script (revised) — Anshuman Pandey

Recommendation:

No action is requested of the UTC.

Comments:

This is a proposal to encode the Zou script (also referred to as Zoulai or Zolai in the document) used for the Zou language used by a community in Northeast India and around. The script was invented in 1952. Despite being adopted by a community organization in 1976 and recognized by the local government in 1991, the proposal does not include any evidence of or reference to any publications written in the script. Two primers are mentioned from the 1990s, but there is no evidence showing them (or any other primer). The section on stability is empty (there is a note of later expansion from 25 to 27 characters, while L2/10-254 indicates the script is cased using 90 characters in 2009) and some figures are not cited or attributed. The data in section 8 does not agree in names or numbers with other images in the proposal (and uses PUA codepoints). Figure 17 shows ligatures and/or conjuncts, but there is no explanation on how they are to be represented. The ISO form indicates no contact with the community and wrong number of characters and does not acknowledge the previous proposal. Without evidence of the script being used and fixing significant deficiencies in the proposal, this cannot be progressed. The SEI might be able to provide some guidance on authoring script proposals.



1.7 Kadu #914

Documents:

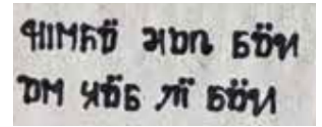
[L2/26-257](#): Proposal for encoding the Kadu script in the UCS — Kadu Than Naing

Recommendation:

No action is requested of the UTC.

Comments:

This is a proposal to encode a recent script for writing the Kadu language of Myanmar, invented between 2019 and 2022. The creator of the script is the author of the proposal, and all the evidence is about the script, there are no examples of any literature or other materials written in the script. The proposal also mentions 4 other scripts invented for the language after 2010, but suggests they were ultimately not adopted.



The SEW views this script to be too new to provide the stability and usage guarantees expected when encoding writing systems. We would like to see evidence of the script being used independently across the community, demonstrating that it succeeded unlike the other ones. This is a reminder that Unicode cannot be used as a tool to establish or enforce new practice on communities, only scripts with already established usage can be considered.

1.8 Wei Rma #862 · #168

Documents:

[L2/26-223](#): Proposal to encode Wei Rma in the UCS — Eiso Chan, et al

Related:

[L2/26-058](#): Proposal to encode Rma script to UCS — Eiso Chan, et al

[L2/26-010](#): Recommendations to UTC #186 (January 2026) on Script Proposals — SEW / Jan Kučera, et al

[L2/25-129](#): Proposal to encode Rma script to UCS — Eiso Chan, et al

[L2/22-130](#): Preliminary proposal to encode Rma script — Eiso Chan, et al

Recommendation:

No action is requested of the UTC.

Comments:

This is a proposal to encode Wei Rma, a script written by the Qiang people in China to record the Qiang language. The script was invented in 2016 and started being used in 2019. An earlier version of the proposal was seen by UTC #186 under the name Rma, where we noted some feedback to the authors. The feedback was addressed, including the revision of the script name and proposed code (Wrma). While the encoding model is still being reviewed, the script is being used commercially and the SEW believes the need for public interchange has been demonstrated. The script was roadmapped at U+16000..U+1604F and it is suggested the UTC seeks feedback on the proposal from ISO/IEC JTC 1/SC 2.



1.9 Wagindara #864 · #483

Documents:

[L2/26-224](#): Proposal to encode Wagindara in the UCS — Yehenara Jiaxu, et al

Related:

[L2/24-228](#): Recommendations to UTC #181 (November 2024) on Script Proposals — Jan Kučera, et al

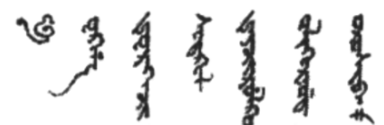
[L2/24-242](#): Preliminary proposal to encode Wagindara letters for the Mongolian script — YEHENARA Jiaxu, et al

Recommendation:

No action is requested of the UTC.

Comments:

The Wagindara script is an alphabetic script developed for the Buryat Mongolian language in the first decade of the 20th century. It was only used for a short time south of Lake Baikal. The authors are arguing the script is separate from Mongolian, and it was roadmapped at U+16050..U+1607F. Discussion focused on whether the evidence meets the SEW expectations for encoding due to its short usage and only one recent publication. The whole corpus consists of a handful of texts. Often with historical scripts, there are efforts underway to digitize and publicly interchange historical documents for research. It is unclear that is the case here, and we would like the authors to update the proposal with some comments on the modern community and need for encoding. Other feedback included to not propose a tatweel-like character unless it has been attested, and to not use a stylized font for the representative glyphs, since we expect them to be very similar to the Mongolian ones.



1.10 Quikscript #923

Documents:

[L2/26-243](#): Proposal to encode Quikscript in the UCS — Karl Pentzlin

Related:

[Preliminary proposal to encode Quikscript characters in the UCS \[not posted\] — Michael Everson](#)

[L2/20-090](#): Comparison of Quikscript vs. Shavian — Deborah Anderson

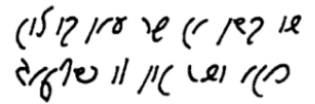
[L2/20-105](#): Recommendations to UTC #163 April 2020 on Script Proposals — Deborah Anderson

Recommendation:

No action is requested of the UTC.

Comments:

Quikscript is a phonographic alphabet for English designed and published in 1966. The evidence points to a community portal, forums and an application, but no printed publications beyond the inventor's manual. There is no information on the community size or stability of the script. Other feedback provided to the author:



- The proposal is bringing in two X characters (EXCITE and EXAM) from the Senior script, but not the half forms or other contractions. Better explanation may be needed about why these two are characters but other features of the Senior script are not.
- Double check the proposed ISO 15924 code as Qskr is not a subsequence of Quikscript.
- Script_Extensions data should be provided for U+00B7 (its current data already includes Shavian.)
- "zhivago" and "oolong" should be proposed as NamesList annotations.

2. New Characters

2.1 Arabic Letter Alef Maksura With Comma Above #907

Documents:

[L2/26-227](#): Proposal to encode ARABIC LETTER ALEF MAKSURA WITH COMMA ABOVE — Takehiko Kyono

Related:

[L2/20-289](#): Request for glyph changes and annotations for Kazakh, Kyrgyz, and Uyghur — Lorna Priest Evans

[L2/11-138](#): Proposal to encode Arabic characters used for Bashkir, Belarusian, Crimean Tatar, Karachay, Karakalpak, and Tatar languages — Ilya Yevlampiey, et al

Recommendation:

1. **[SEW-189-C5] Consensus** : Provisionally assign 1 code point U+0892 ARABIC LETTER ALEF MAKSURA WITH COMMA ABOVE in the Arabic Extended-B block, as described in L2/26-227. [Ref: 2.1 in L2/26-239]
2. **[SEW-189-A8] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 1 code point U+0892 ARABIC LETTER ALEF MAKSURA WITH COMMA ABOVE, as described in L2/26-227. [Ref: 2.1 in L2/26-239]

Comments:

This is a proposal to encode ARABIC LETTER ALEF MAKSURA WITH COMMA ABOVE, used in the writing systems of Tatar and Bashkir, in Arabic-script Nogai and in modern publications on Tatar language published in China. L2/11-201 suggested the character is representable by the sequence of ALEF MAKSURA + DAMMA, but L2/20-289 showed that the mark on U+06C7 ARABIC LETTER U, also used in Turkic languages, is a comma above rather than *damma*. The proposed character is an alif *maqsura* with the same mark, used for recording a vowel specific to Turkic languages. The document provides ample evidence.



2.2 Devanagari Sign Final Anusvara #804

Documents:

[L2/26-206](#): Revised Proposal to encode the Devanagari Sign Final Anusvara — Biswajit Mandal

Related:

[L2/26-181](#): Proposal to encode the Devanagari Sign Final Anusvara in Unicode — Biswajit Mandal

[L2/26-158](#): Recommendations to UTC #188 (July 2026) on Script Proposals — SEW / Jan Kučera, et al

Recommendation:

1. **[SEW-189-C6] Consensus** : Provisionally assign 1 code point U+11B0D DEVANAGARI SIGN FINAL ANUSVARA in the Devanagari Extended-A block, as described in L2/26-206. [Ref: 2.2 in L2/26-239]
2. **[SEW-189-A9] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 1 code point U+11B0D DEVANAGARI SIGN FINAL ANUSVARA, as described in L2/26-206. [Ref: 2.2 in L2/26-239]

Comments:

This document proposes final anusvara for Devanagari, similar to U+11448 NEWA SIGN FINAL ANUSVARA, as a standalone letter used in Jain manuscripts. In the previous recommendations, the SEW suggested that a standalone letter would not serve users well and preferred a combining mark instead, following the Bengali model. However, the proposal author updated the proposal with further evidence demonstrating standalone usage of the character, convincing the SEW to recommend the character as proposed.

ॐ

2.3 Three new characters of Devanagari #913

Documents:

[L2/26-248](#): Proposal to encode three new characters of Devanagari — Biswajit Mandal

Related:

[L2/08-294](#): Placement of characters in Vedic, Devanagari, and Devanagari Extended blocks — Peter Scharf

Recommendation:

1. **[SEW-189-C7] Consensus** : Provisionally assign 1 code point [U+11B0E](#) DEVANAGARI CONTINUATION SIGN in the Devanagari Extended-A block, described in L2/26-248, as amended. [Ref: 2.3 in L2/26-239]
2. **[SEW-189-A10] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 1 code point [U+11B0E](#) DEVANAGARI CONTINUATION SIGN, described in L2/26-248, as amended. [Ref: 2.3 in L2/26-239]
3. **[SEW-189-A11] Action Item** for Robin Leroy, PAG: Consider updating UAX #14 section 5.4 listing all characters that can be used to render a soft hyphen. [Ref: 2.3 in L2/26-239]

Comments:

In this proposal, 3 unrelated Devanagari characters used in Jain manuscripts are requested:

1. DEVANAGARI SIGN JAIN PUSHPIKA. The SEW is not convinced this should be disunified from U+A8F8 DEVANAGARI SIGN PUSHPIKA. L2/08-294, which discusses the U+A8F8 DEVANAGARI SIGN PUSHPIKA at length prior to its encoding, argues for its placement in the Devanagari Extended block rather than Vedic Extensions block specifically due to its usage and presence in Jain manuscripts. We would need some justification why this is not just a stylistic variant of the existing character. It was also noted that in some manuscripts, the character is confusable with U+091B DEVANAGARI LETTER CHA.
2. DEVANAGARI SEPARATOR. This is an apostrophe-like character to mark individual words when text is written in Devanagari without spaces between words. The evidence is rather mixed on whether the line is in between letters interrupting the headline (in which case an existing character like U+0027 APOSTROPHE might be satisfactory), or whether it is above letters. It should be also noted that in Sanskrit, joining words are subject to phonological changes at word boundaries, often resulting in one cluster representing phonemes from both words. Some discussion should be provided on such cases, as the proposed character does not seem to serve them. Finally, this appears to be an editorial annotation over existing text, so a justification why this needs to be encoded in plain text would be desirable. The need for encoding this mark is also questionable as modern digitisations of manuscripts might employ different visualization of word boundaries.
3. DEVANAGARI CONTINUATION SIGN. This is a sign to be written at the end of the lines to indicate when a word (or sentence when word spacing is not employed) is split across lines, similar to how hyphen is used to indicate split words in English and other languages. This seems to be the exact equivalent of U+111DD SHARADA CONTINUATION SIGN, evidence was provided, so SEW supports encoding this character.

ॐ ' ॐ

2.4 Devanagari Correction Mark For Vowel Signs #908

Documents:

[L2/26-221](#): Proposal to encode the Devanagari Correction Mark for Vowel Signs — Biswajit Mandal

Related:

[L2/26-180](#): Proposal to encode Phonetic Correction Mark in Devanagari — Sridatta A, Srinidhi A

Recommendation:

1. **[SEW-189-C8] Consensus** : Provisionally assign 1 code point [U+11B0F](#) DEVANAGARI CORRECTION MARK FOR VOWEL SIGNS in the Devanagari Extended-A block, as described in L2/26-221. [Ref: 2.4 in L2/26-239]
2. **[SEW-189-A12] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 1 code point [U+11B0F](#) DEVANAGARI CORRECTION MARK FOR VOWEL SIGNS, as described in L2/26-221. [Ref: 2.4 in L2/26-239]

Comments:

This is a proposal for a correction mark in Devanagari. It is an above mark found in Jain manuscripts to mark where the right-side U+093E DEVANAGARI VOWEL SIGN AA was inadvertently omitted. This raises some difficult questions about its properties, since it is an above mark in lieu of a right side mark which is sometimes part of a two-part mark. This means that it can go both on consonants as well as on vowel signs. The SEW discussed where it belongs within a cluster, whether atomic encoding of alternative two-part marks would be appropriate, whether it is plain text to begin with, and whether the identity and name of the character are too narrow or too long. It was also compared to the DEVANAGARI PHONETIC CORRECTION MARK proposed in L2/26-180 that indicates intended reading of consonants.



Ultimately the discussion converged on the fact that this is an editorial mark inserted after the whole cluster is written. As such, it should not be interacting with other marks and it should be treated as a top syllable modifier. The proposal originally suggested *bindu* and *candrabindu* should be rendered on top of it, but that is anachronistic and not backed up by evidence. This has now been removed and the proposal has been updated to use the agreed name DEVANAGARI CORRECTION MARK FOR VOWEL SIGNS.

2.5 Dotless I/i with descender #895 · #885

Documents:

[L2/26-176R](#): (Dotless) I/i with descender — Ismael Robles Haloui

Related:

[L2/26-158](#): Recommendations to UTC #188 (July 2026) on Script Proposals — SEW / Jan Kučera, et al

[L2/26-144](#): Unicode request for two Latin orthographic pairs derived from "I i" — Ismael Robles Haloui

Recommendation:

1. **[SEW-189-C9] Consensus** : Provisionally assign 2 code points U+1DFA4 LATIN CAPITAL LETTER DOTLESS I WITH DESCENDER and U+1DFA5 LATIN SMALL LETTER DOTLESS I WITH DESCENDER in the Latin Extended-G block, as described in L2/26-176R. [Ref: 2.5 in L2/26-239]
2. **[SEW-189-A13] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 2 code points U+1DFA4 LATIN CAPITAL LETTER DOTLESS I WITH DESCENDER and U+1DFA5 LATIN SMALL LETTER DOTLESS I WITH DESCENDER, as described in L2/26-176R. [Ref: 2.5 in L2/26-239]

Comments:

This is a proposal for a casing pair of Latin letter I with descender, used in a historical alphabet for the Khakas language for Siberia, with evidence similar to the Sakha letters in the Latin Extended-E block. Both are Turkic languages and in geographical proximity. The proposed casing pair seems to be the only letters missing to record the Khakas language (the remaining ones are in Latin Extended-D). In our last recommendations, we requested few changes including moving the codepoints out of the BMP. The changes were addressed and the SEW is ready to recommend the characters for encoding.



2.6 Latin Capital Turned W #897

Documents:

[L2/26-246](#): New Proposal to add LATIN CAPITAL TURNED W — Denis Moyogo Jacquerye

Related:

[L2/23-134](#): Revised proposal for the addition of LATIN CAPITAL LETTER TURNED W (WG2 N5224) — Michael Everson

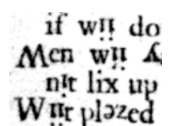
[L2/19-213](#): Proposal for the addition of LATIN CAPITAL LETTER TURNED W (WG2 N5070) — Michael Everson

Recommendation:

1. **[SEW-189-C10] Consensus** : Provisionally assign 1 code point U+AB6F LATIN CAPITAL LETTER TURNED W in the Latin Extended-E block, as described in L2/26-246. [Ref: 2.6 in L2/26-239]
2. **[SEW-189-A14] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 1 code point U+AB6F LATIN CAPITAL LETTER TURNED W as described in L2/26-246. [Ref: 2.6 in L2/26-239]

Comments:

This is a proposal for LATIN CAPITAL LETTER TURNED W. We had had proposals for this character in the past, but the evidence of usage was sparse and limited to works published by the author of the previous proposals. This document provides new evidence from Italian sources where the letter is used as an abbreviation for *abbasso* (down with), as well as in an early 18th century phonetic orthography. Hence the SEW is happy to recommend encoding the character. The properties of the existing U+028D LATIN SMALL LETTER TURNED W should be updated to case-pair with the new character.



2.7 Latin Small Letter Turned Dotless I #905

Documents:

[L2/26-225](#): Unicode request for turned dotless i — Kirk Miller

Related:

[L2/26-154](#): UTC #188 properties feedback & recommendations — PAG / Markus Scherer, Josh Hadley

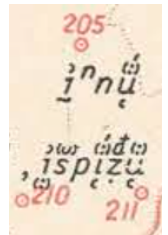
[L2/26-079](#): Unicode request for basic symbols of the Romance dialectological atlases — Kirk Miller

Recommendation:

1. **[SEW-189-C11] Consensus** : Provisionally assign 1 code point U+1DFA6 LATIN SMALL LETTER TURNED DOTLESS I in the Latin Extended-G block, as described in L2/26-225. [Ref: 2.7 in L2/26-239]
2. **[SEW-189-A15] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 1 code point U+1DFA6 LATIN SMALL LETTER TURNED DOTLESS I, as described in L2/26-225. [Ref: 2.7 in L2/26-239]

Comments:

L2/26-079 noted that U+1D09 LATIN SMALL LETTER TURNED I and U+1D4E MODIFIER LETTER SMALL TURNED I, used in Spanish dialectology, were soft-dotted and requested that their properties be changed accordingly. On recommendation from PAG, the UTC concluded that both the existing `Soft_Dotted` property as well as a potential new `Soft_Dotted_Below` property would be too disruptive to the existing systems compared to how rare the characters in question are, and suggested encoding separate dotless characters. The SEW previously expressed an opinion that the convention driving these discussions was a higher-level protocol, but it would be willing to encode the basic building blocks to enable such systems to be created. Therefore, the SEW is recommending to propose a dotless version of the base character. The rendering rules for the rest of the system require further research.



2.8 Ellis's Phonotypic Alphabet for Yahgan and Ona #886

Documents:

[L2/26-245](#): Proposal to encode in the UCS extensions to Ellis's Phonotypic Alphabet for writing Yahgan and Ona — Yoram Meroz, Karl Pentzlin

Related:

[L2/24-277](#): Fifth Revised Proposal to encode characters for the English Phonotypic Alphabet (EPA) in the UCS — Karl Pentzlin

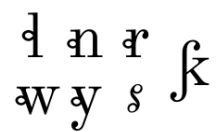
Recommendation:

1. **[SEW-189-C12] Consensus** : Provisionally assign 8 code points: U+1AF4 in the Combining Diacritical Marks Extended block and U+1DFA7..U+1DFAD in the Latin Extended-G block, as described in L2/26-245. [Ref: 2.8 in L2/26-239]
2. **[SEW-189-A16] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 8 code points U+1AF4, U+1DFA7..U+1DFAD, as described in L2/26-245. [Ref: 2.8 in L2/26-239]

Comments:

This is a proposal to include several characters extending the Ellis Phonotypic Alphabet provisionally assigned by 182-C7. The proposed characters were used in the 19th century in the documentation of the Yahgan language, as well as the Ona language. This document underwent several iterations addressing feedback from the SEW.

Notably, SEW did not support encoding uppercase variants of the letters, as they were only attested in charts and no actual text. A character representing *tsh*, *ths* or *sth* was unified with U+01BE LATIN LETTER INVERTED GLOTTAL STOP WITH STROKE. The identity of U+1DFAD LATIN SMALL LETTER ESH-K was questioned, suggesting either f-k or j-k ligature instead, but the phonology does not support that. All the feedback was addressed and the SEW reached consensus to recommend the proposal.



2.9 South Slavic Cyrillic modifier letters #926

Documents:

[L2/26-244](#): Unicode request for South Slavic Cyrillic modifier letters — Kirk Miller

Related:

[L2/22-010](#): Addendum II to L2/21-107, Cyrillic modifier letters — Kirk Miller

[L2/21-142](#): Addendum to L2/21-107, Cyrillic modifier letters — Kirk Miller

[L2/21-107](#): Unicode request for Cyrillic modifier letters — Kirk Miller

Recommendation:

1. **[SEW-189-C13] Consensus** : Provisionally assign 6 code points U+1E06E..U+1E073 in the Cyrillic Extended-D block, as described in L2/26-244. [Ref: 2.9 in L2/26-239]
2. **[SEW-189-A17] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 6 code points U+1E06E..U+1E073, as described in L2/26-244. [Ref: 2.9 in L2/26-239]

Comments:

In the cited previous proposals, several Cyrillic modifier letters were recommended, forming the Cyrillic Extended-D block. There are gaps in coverage, particularly for South Slavic languages, due to the lack of evidence of usage. This document provides evidence of 6 of the missing modifier letters, which are hereby recommended for encoding.

Љ Њ Ћ Ц Ъ Ь

2.10 Myanmar letters for Kham Dom Pa'O #867

Documents:

[L2/26-222](#): Proposal to encode Myanmar letters for Kham Dom Pa'O — Ben Mitchell

Related:

[ID20260718022511](#): Accumulated Feedback on PRI #544

Recommendation:

No action is requested of the UTC.

Comments:

This is a proposal to encode additional Myanmar letters used in Kham Dom Pa'O. We had already received public feedback about these letters as missing. The SEW is generally in support of the proposal, but requested some changes to character names, and review of collation is still in progress. Feedback shared with the authors include:



- Use names consistent with existing names in the block (English naming, SOUTHERN PAO; native names can be proposed as annotations)
- Confusability and co-occurrence with the tone mark (or lack thereof) should be discussed
- Clarify the highlighting of proposed characters in section 2 (there are more highlighted rows than proposed characters)
- There is a typo in a codepoint for laPan when discussing htuaamPla in section 3.
- Provide examples of how words should sort, before section 4.
- Add header (and optional description) for NamesList.txt for the proposed characters.

2.11 Malayalam Letter Vedic Double Anusvara #921

Documents:

[L2/26-251](#): Proposal to encode Vedic Double Anusvara in Malayalam — Sridatta A, Srinidhi A

Recommendation:

No action is requested of the UTC.

Comments:

This is a proposal to encode double anusvara for Malayalam. The SEW does not have any objections to the character, but suggests that adding Malayalam to the scripts using U+1135F GRANTHA LETTER VEDIC DOUBLE ANUSVARA might be an appropriate alternative solution given the history of the script. The authors should discuss the advantages and disadvantages of using the Grantha letter.



2.12 Turkestan Brāhmī #932

Documents:

[L2/26-252](#): Preliminary proposal for Turkestan Brahmi within the Brahmi block — Francesco Catoni

Related:

[L2/15-236](#): Proposal to Encode the Tocharian Script (supersedes L2/15-023) — Lee Wilson

[L2/15-023](#): Preliminary Proposal to Encode the Tocharian Script — Lee Wilson

[L2/15-022](#): Preliminary Proposal to Encode the Khotanese Script — Lee Wilson

[L2/14-192](#): Preliminary Proposal to Encode the Turkestani Script — Lee Wilson

Recommendation:

No action is requested of the UTC.

Comments:

This is a proposal to extend the existing Brahmi block with characters needed for North Turkestan Brahmi and South Turkestan Brahmi. The SEW was in agreement that the characters are not what Unicode classifies as Brahmi, and if these were to be encoded, it might be more appropriate as a new script or even two separate scripts. We asked the authors to provide a comparison across the two varieties and Brahmi.



2.13 Arabic Characters from Indo-Pak Mushafs #924 · #767

Documents:

[L2/26-229](#): Proposal to Encode Arabic Characters from Indo-Pak Mushafs — Rikza F. Sh.

Related:

[L2/26-100](#): Recommendations to UTC #187 (April 2026) on Script Proposals — SEW / Jan Kučera, et al

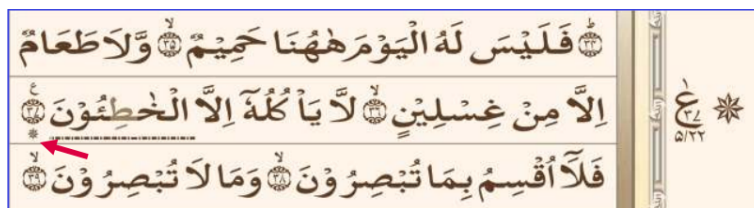
[L2/26-125](#): Proposal to Encode Arabic Characters from Indo-Pak Mushafs — Rikza F. Sh.

Recommendation:

No action is requested of the UTC.

Comments:

This proposal presents a set of Arabic characters used in typographic traditions found in specialized Quranic manuscripts and printed mushafs developed in the Indian subcontinent. We provided extensive feedback in L2/26-100 and most of it was addressed.



Some of the proposed characters (TAKHALLUS BELOW, PLACE OF DUA ABOVE, PLACE OF DUA BELOW, and IHTIYAAT

BELOW) may need to be represented using a higher-level protocol. Another character, the ARABIC SMALL HIGH WAQF LAZIM is being proposed as a combining mark, but there is no base character in the evidence - on the contrary, it carries a mark above it. In order to progress the proposal, we suggest the above mentioned characters be removed or split into a separate proposal. This will require shifting the remaining characters into a continuous block. Finally, the names of U+10E48 ARABIC OPEN SMALL DOTLESS HEAD OF KHAH and U+10E49 ARABIC ROTATED OPEN SMALL DOTLESS HEAD OF KHAH should include the word HIGH after U+06E1 ARABIC SMALL HIGH DOTLESS HEAD OF KHAH.

2.14 Benjamin Franklin's reformed alphabet #925

Documents:

[L2/26-230](#): Proposal to encode four Latin letters of Benjamin Franklin's reformed alphabet — Karl Pentzlin

Related:

[L2/22-248](#): Recommendations to UTC #173 October 2022 on Script Proposals — Deborah Anderson, et al

[L2/22-215](#): Proposal to encode missing Latin letters from Benjamin Franklin's phonetic alphabet — Nikita Manulov

WG2 [N3555](#): Exploratory proposal to encode Germanicist, Nordicist, and other phonetic characters — Michael Everson

Recommendation:

No action is requested of the UTC.

Comments:

Benjamin Franklin's reformed English alphabet introduced six new lowercase letters. Two of them are representable using U+0265 LATIN SMALL LETTER TURNED H and U+014B LATIN SMALL LETTER ENG. This is a proposal for the remaining 4. The SEW recognizes that this is a small corpus with letters that have not been established in later usage, but it is an important corpus with demonstrable need for interchange.



A previous proposal in 2022 resulted in requests for better evidence, glyphs, naming, community identification, representability analysis and SMP allocation, but there was no follow-up from the author. The current document is trying to address the feedback including stronger evidence.

The SEW is generally in favor of the proposal which now uses functional names avoiding issues with handwritten forms being different, but provided feedback focusing mostly on the proposed representative glyphs. We are hoping the author will be able to provide more and better quality examples especially and include a comparison table of the attested variants.

3. New Symbols

3.1 Leibniz: 6 letterlike symbols #691

Documents:

- [L2/26-259](#): Request to re-order 5 letterlike symbols in the 1DB00 block — Andreas Stötzner
- [L2/26-142](#): Proposal to encode 6 letterlike symbols — Uwe Mayer, et al

Recommendation:

- [SEW-189-C14] Consensus** : Provisionally assign 5 code points: 1DB1D ALPHA X SYMBOL, 1DB1E GREEK OMICRON UPSILON SYMBOL, 1DB1F DOUBLE SMALL SIGMA SYMBOL, 1DB20 FUNCTIO DIFFERENTIATA SYMBOL, 1DB21 DOUBLE X SYMBOL in the Miscellaneous Symbols and Arrows Extended block, described in L2/26-142, ordered in L2/26-259, as amended. [Ref: 3.1 in L2/26-239]
- [SEW-189-A18] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 5 code points U+1DB1D..U+1DB21, described in L2/26-142, ordered in L2/26-259, as amended. [Ref: 3.1 in L2/26-239]
- [SEW-189-A19] Action Item** for Debbie Anderson, SAH: Work with the proposal authors to submit a compliant font for Leibniz letterlike symbols to Unicode. [Ref: 3.1 in L2/26-239]
- [SEW-189-A20] Action Item** for Josh Hadley, Roozbeh Pournader, PAG: Add confusable information for U+1DB1E GREEK OMICRON UPSILON SYMBOL and U+0223 LATIN SMALL LETTER OU, for a future version of the standard. [Ref: 3.1 in L2/26-239]

Comments:

This proposal from the Leibniz project group for 6 letterlike symbols was recommended for provisional assignment in the previous report, but the UTC remanded it back over the questions raised about the name of OMICRON UPSILON SYMBOL. The SEW carefully reviewed the name and recommends changing it to GREEK OMICRON UPSILON SYMBOL, following U+03D0 GREEK BETA SYMBOL and others. SEW is recommending provisionally assigning five characters; the sixth symbol was already accepted by 188-C17 as a variation sequence.



3.2 Leibniz: Miscellaneous symbols #692

Documents:

- [L2/26-140R](#): Proposal to encode miscellaneous mathematical symbols — Uwe Mayer, et al

Recommendation:

- [SEW-189-C15] Consensus** : Provisionally assign 10 code points U+1DB22..U+1DB2B in the Miscellaneous Symbols and Arrows Extended block, as described in L2/26-140R. [Ref: 3.2 in L2/26-239]
- [SEW-189-A21] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 10 code points U+1DB22..U+1DB2B, as described in L2/26-140R. [Ref: 3.2 in L2/26-239]
- [SEW-189-A22] Action Item** for Debbie Anderson, SAH: Work with the proposal authors to submit a compliant font for Leibniz miscellaneous symbols to Unicode. [Ref: 3.2 in L2/26-239]

Comments:

This proposal from the Leibniz project group for 10 miscellaneous symbols was recommended for provisional assignment in the previous report, but the UTC remanded it back over the questions raised about the distinction between U+1DB24 PROPORTION WITH ONE STROKE and U+22A4 DOWN TACK. The proposal has now been updated providing justification for the disunification, notably the need for the horizontal bar to be positioned at the math axis height, and greater width than U+22A4 DOWN TACK.



3.3 Turned Digit One for Balanced Ternary #915

Documents:

- [L2/26-228](#): Proposal to encode TURNED DIGIT ONE for Balanced Ternary — Adam Victor

Recommendation:

No action is requested of the UTC.

Comments:

This document proposes to encode the symbol TURNED DIGIT ONE for balanced ternary systems. However, the proposal neither provides any evidence that such symbol is used by the community, nor does it demonstrate any need for public interchange. One publication from 1963 is mentioned by reference. The modern community appears to be using either T or -1. The role of Unicode is not to establish new orthography practices. Without sufficient attestation and demonstrated need for interchange, this character cannot be recommended.



3.4 Big5 ETen Extension Box Drawing Characters #911

Documents:

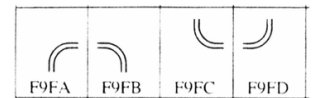
[L2/26-220](#): Proposal to Encode 4 Unencoded Big5 ETen Extension Box Drawing Characters — Night Koo, Luke036

Recommendation:

No action is requested of the UTC.

Comments:

This is a proposal to encode 4 box drawing characters of double arc quadrants in Big5 ETen Extension. The SEW does not have objections to encoding these characters in principle, but did not find enough justification for encoding. The authors were asked to update the proposal with modern usage, the need for public interchange and reasons why the existing 25-year-old mapping is insufficient. Other feedback included leaving the obsolete `Unicode_1_Name` field empty and avoiding suggesting the mapping tables are official, because they are not.



4. Changes to Characters

4.1 Report on UTC 158-A106 #931

Documents:

[L2/26-255](#): Some Changes to Indic_Syllabic_Category: A Report on UTC 158-A106 — Roozbeh Pournader

Recommendation:

1. **[SEW-189-C16] Consensus** : Update the Indic syllabic category of U+0ECC, U+1CE9..U+1CEC, U+1CEE..U+1CF0, U+1CFA, U+A8F4..U+A8F7, U+11448 and U+114C5 as described in L2/26-255, for Unicode 19.0. [Ref: 4.1 in L2/26-239]
2. **[SEW-189-A23] Action Item** for Roozbeh Pournader, PAG: Update the Indic syllabic category of U+0ECC, U+1CE9..U+1CEC, U+1CEE..U+1CF0, U+1CFA, U+A8F4..U+A8F7, U+11448, U+114C5 and the comment on Bindu entries, as described in L2/26-255, for Unicode 19.0. [Ref: 4.1 in L2/26-239]
3. **[SEW-189-A24] Action Item** for Ken Whistler, Charts: Consider adding annotation to U+0ECC LAO CANCELLATION MARK, as described in L2/26-255. [Ref: 4.1 in L2/26-239]

Comments:

This is a report on research requested by 158-A106. It suggests assigning Indic syllabic category value of `Bindu` to several characters, including U+1CFA VEDIC SIGN DOUBLE ANUSVARA ANTARGOMUKHA which is currently `Consonant_Placeholder`, and adding annotations to U+0ECC LAO CANCELLATION MARK, and changing its Indic syllabic category to `Consonant_Killer`. The comment on Bindu entries in the data file should be updated accordingly. The SEW agrees with these conclusions and thanks Roozbeh Pournader for his work on this action item, which has thus been resolved.

4.2 Report on UTC 164-A63 #933

Documents:

[L2/26-256](#): Categorizing Some Uncategorized Indic Characters: A Report on UTC 164-A63 — Roozbeh Pournader

Recommendation:

1. **[SEW-189-C17] Consensus** : Update the Indic syllabic category of U+1CE2..U+1CE8 and U+1CED, and Indic positional category of U+1CF8..U+1CF9 as described in L2/26-256, for Unicode 19.0. [Ref: 4.2 in L2/26-239]
2. **[SEW-189-A25] Action Item** for Roozbeh Pournader, PAG: Update the Indic syllabic category of U+1CE2..U+1CE8 and U+1CED, and Indic positional category of U+1CF8..U+1CF9 as described in L2/26-256, for Unicode 19.0. [Ref: 4.2 in L2/26-239]

Comments:

This is a report on research requested by 164-A63. It suggests assigning Indic syllabic category value of `Cantillation_Mark` to several Vedic signs and Indic positional category value of `Top` to U+1CF8 VEDIC TONE RING ABOVE and U+1CF9 VEDIC TONE DOUBLE RING ABOVE. The SEW agrees with these conclusions and thanks Roozbeh Pournader for his work on this action item, which has thus been resolved.

4.3 Character properties of U+0CD5 KANNADA LENGTH MARK #918

Documents:

[L2/26-250](#): Character properties of U+0CD5 KANNADA LENGTH MARK — Sridatta A, Srinidhi A

Related:

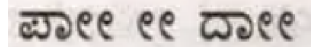
[L2/19-079](#): Proposal for revision of Unicode Bangla Block — ICT Div, Gov't of Bangladesh

Recommendation:

No action is requested of the UTC.

Comments:

In this document, authors are seeking advice on how to use vowel signs in isolation. They are suggesting changing the general category and line breaking properties of a rather common character, U+0CD5 KANNADA LENGTH MARK, which would be considerably disruptive to various implementations. Moreover, canonically equivalent strings would exhibit different line breaking. The SEW cannot support such change. However, it recognizes the demonstrated usage of this and similar other characters which came up in the past. There is an outstanding action item 159-A79 to figure out how to support these in general, originally triggered by similar situation described in L2/19-079.



5. Other

5.1 Vithkuqi annotations #853

Documents:

[L2/26-208](#): Unicode request for Vithkuqi annotations II — Francesco Catoni, Kirk Miller

Related:

[L2/20-187](#): Proposal for encoding the Vithkuqi script in the SMP (revised, WG2 N5138R2) — SEI / Michael Everson

Recommendation:

1. **[SEW-189-A26] Action Item** for Ken Whistler, Charts: Consider removing annotations for characters in the Vithkuqi block, as described in L2/26-208. [Ref: 5.1 in L2/26-239]

Comments:

This is a request to remove "used in 19th century orthography" annotations from characters in the Vithkuqi block, since the entire script was used exclusively in the 19th century. The proposal is also requesting marking some codepoints as reserved. These are not listed in the names list because they are at the end of sections, which can be misunderstood as being available while they are reserved for specific characters. This is likely a charts production limitation and we defer the request to the Charts group.

5.2 Report on UTC 159-A81 #927

Documents:

[L2/26-247](#): Commas and Cedillas: Report on UTC 159-A81 — Roozbeh Pournader

Recommendation:

1. **[SEW-189-A27] Action Item** for Roozbeh Pournader, EDC: Update the recommendations on the default rendering of U+0327 COMBINING CEDILLA in chapter 7 of the Core Specification, based on the descriptions in L2/26-247. [Ref: 5.2 in L2/26-239]

Comments:

This is a report on research requested by 159-A81 on feedback submitted by David Corbett regarding preferred rendering of cedilla versus comma below in the Core Specification. The report suggests accepting the feedback and making changes to the Core Specification, most notably not recommending comma glyph below t as the default rendering behavior. The SEW agrees with these conclusions and thanks Roozbeh Pournader for his work on this action item, which has thus been resolved.

5.3 On U+055F ARMENIAN ABBREVIATION MARK #928 · #131

Documents:

[L2/26-213](#): On U+055F ARMENIAN ABBREVIATION MARK — Roozbeh Pournader

Recommendation:

No action is requested of the UTC.

Comments:

This is an informational report on the existing usage of U+055F ARMENIAN ABBREVIATION MARK intended for future discussions about the character. No action is requested at this time.

5.4 Repha in Kannada #922

Documents:

[L2/26-253](#): Repha and its behaviour in Kannada script — Sridatta A, Srinidhi A

Related:

[L2/09-372](#): Proposal to encode the Grantha script — Shriramana Sharma

Recommendation:

No action is requested of the UTC.

Comments:

This document discusses two issues regarding the behavior of repha in Kannada.

1. The sequence of RA + VIRAMA + consonant in Kannada is rendered using repha on the right side of the consonant, unless the consonant is RA. The sequence RA + VIRAMA + RA (rra) should be rendered using subjoined RA. Not all fonts on all platforms seem to be doing that, so the document is suggesting adding appropriate language to the Core Specification. Before recommending such addition, we would like to see some survey of existing fonts and their behaviour on various platforms.
2. The U+0CD5 KANNADA LENGTH MARK is a spacing right-side mark. When in a cluster with repha, there are two possible orders of the two right-side marks. The document is seeking feedback from SEW and other experts on how this variation needs to be represented and rendered. The question from SEW is whether there is any semantic difference between the two orders.

5.5 Nukta in Sinhala #950

Documents:

[L2/26-254](#): Nukta in Sinhala script — Sridatta A, Srinidhi A

Recommendation:

No action is requested of the UTC.

Comments:

This is a document showing the use of *nukta* in Sinhalese. The Core Specification currently states *"Some letters, such as U+0DBB SINHALA LETTER RAYANNA and U+0DB1 SINHALA LETTER DANTAJA NAYANNA, may be modified by adding the equivalent of a nukta. There is, however, no nukta presently encoded in the Sinhala block."* The document contains enough evidence for SEW to recommend a new character. It is posted to L2 registry to solicit feedback from interested parties.

5.6 Cross-Writing-System Ambiguity in Mongolian Script #904

Documents:

[L2/26-207](#): Resolving Cross-Writing-System Ambiguity in Mongolian Script — Kushim Jiang

Related:

[L2/17-115R](#): Glyph corrections for Ali Gali final forms in the Mongolian block (revised) — Weizhe Zheng, et al

Recommendation:

No action is requested of the UTC.

Comments:

This is a document seeking advice from SEW in regards to Mongolian characters used for writing systems or orthographic traditions such as Hudum, Todo, Sibe or Manchu, in which some positional variants vary or do not occur at all, but font designers still need to supply a glyph. For different glyph shapes based on language, including positional variants, language tags are expected to be utilized. The rest of the questions are all down to the fact that in some languages, some positional variants do not occur. The document suggests rendering them either visibly broken, or using glyphs synthesized from other positional variants. The SEW reached consensus that this is unnecessary, if not undesirable. If certain positional variants do not occur in some languages, then it does not matter what is rendered. Specifically, the positional variant of a language for which it is defined can be used.

It was noted that while some final forms might not be occurring, during typing, to-be medial forms would be appearing as final forms until the subsequent characters are typed. While this might be confusing to users, it was deemed acceptable to show unexpected final forms during typing. Special input methods can be developed to alleviate this behavior, for example using the LDML keyboard layouts.

6. In Process

Documents:

[L2/26-195](#): Proposal to Encode 4 Old Sundanese Characters — Febri Muhammad Nasrullah

[L2/26-146](#): Proposal to Encode Maya Hieroglyphs Extended-A — Dr. Gabrielle Vail, et al

[L2/26-249](#): Proposal to encode Devanagari Double Separator — Biswajit Mandal

Appendix

List of documents covered by this report:

[L2/26-055R](#): Proposal to Encode the Ndiko Jonam Script in the UCS — Oreen Yousuf
[L2/26-087](#): Proposal for Encoding the N'ti Script in the UCS (supercedes L2/26-028R) — Oreen Yousuf, Ibrahima Ceesay
[L2/26-132R](#): Proposal for the encoding of « KORE SEBELI » — Mohamed Bentoura Bangoura
[L2/26-140R](#): Proposal to encode miscellaneous mathematical symbols — Uwe Mayer, et al
[L2/26-142](#): Proposal to encode 6 letterlike symbols — Uwe Mayer, et al
[L2/26-145R3](#): Proposal to Encode Codical Maya Hieroglyphs — Dr. Des. Carlos Pallan, et al
[L2/26-146](#): Proposal to Encode Maya Hieroglyphs Extended-A — Dr. Gabrielle Vail, et al
[L2/26-176R](#): (Dotless) I/i with descender — Ismael Robles Haloui
[L2/26-195](#): Proposal to Encode 4 Old Sundanese Characters — Febri Muhammad Nasrullah
[L2/26-206](#): Revised Proposal to encode the Devanagari Sign Final Anusvara — Biswajit Mandal
[L2/26-207](#): Resolving Cross-Writing-System Ambiguity in Mongolian Script — Kushim Jiang
[L2/26-208](#): Unicode request for Vithkuqi annotations II — Francesco Catoni, Kirk Miller
[L2/26-213](#): On U+055F ARMENIAN ABBREVIATION MARK — Roozbeh Pournader
[L2/26-220](#): Proposal to Encode 4 Unencoded Big5 ETen Extension Box Drawing Characters — Night Koo, Luke036
[L2/26-221](#): Proposal to encode the Devanagari Correction Mark for Vowel Signs — Biswajit Mandal
[L2/26-222](#): Proposal to encode Myanmar letters for Kham Dom Pa'O — Ben Mitchell
[L2/26-223](#): Proposal to encode Wei Rma in the UCS — Eiso Chan, et al
[L2/26-224](#): Proposal to encode Wagindara in the UCS — Yehenara Jiayu, et al
[L2/26-225](#): Unicode request for turned dotless i — Kirk Miller
[L2/26-227](#): Proposal to encode ARABIC LETTER ALEF MAKSURA WITH COMMA ABOVE — Takehiko Kyono
[L2/26-228](#): Proposal to encode TURNED DIGIT ONE for Balanced Ternary — Adam Victor
[L2/26-229](#): Proposal to Encode Arabic Characters from Indo-Pak Mushafs — Rikza F. Sh.
[L2/26-230](#): Proposal to encode four Latin letters of Benjamin Franklin's reformed alphabet — Karl Pentzlin
[L2/26-243](#): Proposal to encode Quikscript in the UCS — Karl Pentzlin
[L2/26-244](#): Unicode request for South Slavic Cyrillic modifier letters — Kirk Miller
[L2/26-245](#): Proposal to encode in the UCS extensions to Ellis's Phonotypic Alphabet for writing Yahgan and Ona — Yoram Meroz, Karl Pentzlin
[L2/26-246](#): New Proposal to add LATIN CAPITAL TURNED W — Denis Moyogo Jacquerye
[L2/26-247](#): Commas and Cedillas: Report on UTC 159-A81 — Roozbeh Pournader
[L2/26-248](#): Proposal to encode three new characters of Devanagari — Biswajit Mandal
[L2/26-249](#): Proposal to encode Devanagari Double Separator — Biswajit Mandal
[L2/26-250](#): Character properties of U+0CD5 KANNADA LENGTH MARK — Sridatta A, Srinidhi A
[L2/26-251](#): Proposal to encode Vedic Double Anusvara in Malayalam — Sridatta A, Srinidhi A
[L2/26-252](#): Preliminary proposal for Turkestan Brahmi within the Brahmi block — Francesco Catoni
[L2/26-253](#): Repha and its behaviour in Kannada script — Sridatta A, Srinidhi A
[L2/26-254](#): Nukta in Sinhala script — Sridatta A, Srinidhi A
[L2/26-255](#): Some Changes to Indic_Syllabic_Category: A Report on UTC 158-A106 — Roozbeh Pournader
[L2/26-256](#): Categorizing Some Uncategorized Indic Characters: A Report on UTC 164-A63 — Roozbeh Pournader
[L2/26-257](#): Proposal for encoding the Kadu script in the UCS — Kadu Than Naing
[L2/26-258](#): Proposal to Encode the Zou Script and Numerals in Unicode — Dr. T Marvin Zou
[L2/26-259](#): Request to re-order 5 letterlike symbols in the 1DB00 block — Andreas Stötzner

List of action items per assignee:

Debbie Anderson, SAH

- Work with the proposal authors to submit a compliant font for N'ti to Unicode.
- Work with the proposal authors to submit a compliant font for Ndiko Jonam to Unicode.
- Work with the proposal authors to submit a compliant font for Koré Sébèli to Unicode.
- Work with the proposal authors to submit a compliant font for Leibniz letterlike symbols to Unicode.
- Work with the proposal authors to submit a compliant font for Leibniz miscellaneous symbols to Unicode.

Josh Hadley, PAG

- Add confusable information for U+1DB1E GREEK OMICRON UPSILON SYMBOL and U+0223 LATIN SMALL LETTER OU, for a future version of the standard.

Ken Whistler, RMG

- Update the Pipeline to include 513 code points U+15500..U+156DC, U+156EC..U+1570F, as described in L2/26-145R3.
- Update the Pipeline to include 47 provisionally assigned code points U+1CB80..U+1CBA4, U+1CBA6..U+1CBAF, in a new Nti block: U+1CB80..U+1CBAF, as described in Section V of L2/26-087.
- Update the Pipeline to include 69 provisionally assigned code points U+1CB30..U+1CB49, U+1CB51..U+1CB71, and U+1CB76..U+1CB7F, in a new Ndiko Jonam block: U+1CB30..U+1CB7F, as described in L2/26-055R.
- Update the Pipeline to include 93 provisionally assigned code points U+15B00..U+15B52, U+15B56..U+15B5F, in a new Kore Sebeli block: U+15B00..U+15B5F, described in L2/26-132R, as amended.
- Update the Pipeline to include 1 code point U+0892 ARABIC LETTER ALEF MAKSURA WITH COMMA ABOVE, as described in L2/26-227.
- Update the Pipeline to include 1 code point U+11B0D DEVANAGARI SIGN FINAL ANUSVARA, as described in L2/26-206.
- Update the Pipeline to include 1 code point U+11B0E DEVANAGARI CONTINUATION SIGN, described in L2/26-248, as amended.
- Update the Pipeline to include 1 code point U+11B0F DEVANAGARI CORRECTION MARK FOR VOWEL SIGNS, as described in L2/26-221.
- Update the Pipeline to include 2 code points U+1DFA4 LATIN CAPITAL LETTER DOTLESS I WITH DESCENDER and U+1DFA5 LATIN SMALL LETTER DOTLESS I WITH DESCENDER, as described in L2/26-176R.
- Update the Pipeline to include 1 code point U+AB6F LATIN CAPITAL LETTER TURNED W as described in L2/26-246.
- Update the Pipeline to include 1 code point U+1DFA6 LATIN SMALL LETTER TURNED DOTLESS I, as described in L2/26-225.
- Update the Pipeline to include 8 code points U+1AF4, U+1DFA7..U+1DFAD, as described in L2/26-245.
- Update the Pipeline to include 6 code points U+1E06E..U+1E073, as described in L2/26-244.
- Update the Pipeline to include 5 code points U+1DB1D..U+1DB21, described in L2/26-142, ordered in L2/26-259, as amended.
- Update the Pipeline to include 10 code points U+1DB22..U+1DB2B, as described in L2/26-140R.

Ken Whistler, Charts

- Consider adding annotation to U+0ECC LAO CANCELLATION MARK, as described in L2/26-255.
- Consider removing annotations for characters in the Vithkuqi block, as described in L2/26-208.

Robin Leroy, PAG

- Consider updating UAX #14 section 5.4 listing all characters that can be used to render a soft hyphen.

Roosbeh Pournader, PAG

- Add confusable information for U+1DB1E GREEK OMICRON UPSILON SYMBOL and U+0223 LATIN SMALL LETTER OU, for a future version of the standard.
- Update the Indic syllabic category of U+0ECC, U+1CE9..U+1CEC, U+1CEE..U+1CF0, U+1CFA, U+A8F4..U+A8F7, U+11448, U+114C5 and the comment on Bindu entries, as described in L2/26-255, for Unicode 19.0.
- Update the Indic syllabic category of U+1CE2..U+1CE8 and U+1CED, and Indic positional category of U+1CF8..U+1CF9 as described in L2/26-256, for Unicode 19.0.

Roosbeh Pournader, EDC

- Update the recommendations on the default rendering of U+0327 COMBINING CEDILLA in chapter 7 of the Core Specification, based on the descriptions in L2/26-247.