

TO: UTC
FROM: Deborah Anderson, Manish Goregaokar, Jan Kučera, Ken Whistler, Roozbeh Pournader, and Peter Constable¹
SUBJECT: Recommendations to UTC #178 January 2024 on Script Proposals
DATE: January 17, 2024

The Script Ad Hoc group met on November 17 and December 22, 2023, and January 5, 2024, in order to review proposals. The following represents feedback on proposals that were available when the group met.

Table of Contents

A. PROPOSALS REQUIRING UTC ACTION	2
I. AFRICA	2
1. Beria Erfe	2
2. Egyptian Hieroglyphs	3
2a Extended Egyptian Hieroglyphs Repertoire	3
2b Egyptian Hieroglyphs Database	4
II. MIDDLE EASTERN SCRIPTS	4
3. Arabic	4
3a Eight More Honorifics	4
3b Potential disunification of KASHMIRI YEH	5
III. EAST ASIAN SCRIPTS	6
4. CJK Punctuation	6
IV. SOUTH AND CENTRAL ASIAN SCRIPTS	7
5. Kirat Rai and Tulu-Tigalari	7
6. Sunuwar	8
7. Vedic Punctuation	8
8. Mongolian	9
V. SOUTHEAST ASIA, INDONESIA, AND OCEANIA	10
9. Batak	10
VI. SYMBOLS	10
10. Numbers with slashes used in figured bass (musical symbols)	10
11. One Chess Symbol	11

¹ Also participating were Basit Ali, Ege Altunsu, Gavin Jared Bala, Craig Cummings, Lorna Evans, Karljürgen Feuerherm, Liang Hai, Ned Holbrook, Anushah Hossain, John Hudson, Richard Ishida, Maria Jarzyna, Kushim Jiang, James Kirby, Robin Leroy, Norbert Lindenberg, Kamal Mansour, Kirk Miller, Cheon Hyeong Sim, Logan Simpson, Michel Suignard, Harald Tveiten, Lawrence Wolf-Sonkin, and Ben Yang. The text for the comments and recommendations was based on notes primarily by Manish Goregaokar and Jan Kučera with assistance from Debbie Anderson.

12. Proposal to Change BidiMirroring property for U+226D	12
13. Shatranj symbols	13
14. Three musical symbols	13
15. Unicode request for Stein-Zimmermann quarter-tone accidentals (musical symbols)	14
VII. FEEDBACK	15
16. Bamum character name error	15
17. On Provisionally Assigned Characters' Names	15
VIII. OTHER	16
18. Additions to ScriptExtensions	16
19. DoNotEmit.txt	17
20. Requirements and Process for Changing Script Status for Identifier Use	18
B. DOCUMENTS NOT REQUIRING UTC ACTION	18
21. Old Hungarian	18
22. Cuneiform	19
23. Ethereum symbol	19
24. Bété	20
C. IN PROCESS	20

A. PROPOSALS REQUIRING UTC ACTION

I. AFRICA

1. Beria Erfe

Document: [L2/24-004](#) Proposal for encoding the Beria Erfe (Zaghawa) script in the SMP of the UCS -- Siddick Adam Issa, et al.

Recommendation: We recommend the UTC make the following disposition:

- **Consensus:** Provisionally assign 50 code points U+16EA0..U+16ED1 in a new Beria Erfe block U+16EA0..U+16EDF for 50 Beria Erfe characters as described in [L2/24-004](#)
- **Action Item** for Uma Umamaheswaran, Roadmap Committee: Update the Roadmap from 3 columns to 4 columns (U+16EA0..U+16EDF), and reflect the name change from "Beria" to "Beria Erfe". [Reference: Section 1 of [L2/24-013](#)]
- **Action Item** for Ken Whistler, UTC: Update the Pipeline to include 50 Beria Erfe characters U+16EA0..U+16ED1 in a new Beria Erfe block U+16EA0..U+16EDF. [Reference: Section 1 of [L2/24-013](#)]
- **Action Item** for Lorna Evans, SAH: Send a Beria Erfe font to Michel Suignard.

[Reference: Section 1 of [L2/24-013](#)]

Comments:

This is a proposal for 50 characters for the Beria Erfe script, used to write the Beria language in Chad, Sudan, and the diaspora. It is a bicameral alphabet, with capital letters typically extending below the baseline. Some characters can take accent marks from the Combining Diacritical Marks block.

The script is used by multiple communities which seem to have alignment on its usage, potentially with some differences on usage of combining marks.

In earlier proposals the script name contained “Giray,” which means writing, but this was removed from the name based on Script Ad Hoc recommendations. “Erfe” refers to the livestock branding marks the script was based on, hence “Beria Erfe” distinguishes the script from the branding marks. “Zaghawa” is an exonym of the language and while not in the script name, the proposal asks that it be mentioned alongside the script name in the code chart / spec chapter.

2. Egyptian Hieroglyphs

2a Extended Egyptian Hieroglyphs Repertoire

Document: [L2/23-181R](#) Encoding proposal for an extended Egyptian Hieroglyphs repertoire -- Michel Suignard et al.

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Approve 3,995 Egyptian Hieroglyph characters U+13460..U+143FA in a new Egyptian Hieroglyphs Extended-A block at U+13460..U+143FF, as described in [L2/23-181R](#) for Unicode 16.0 [Reference: Section 2a of [L2/24-013](#)]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline with 3,995 Egyptian Hieroglyph characters U+13460..U+143FA in a new Egyptian Hieroglyphs Extended-A block at U+13460..U+143FF as described in [L2/23-181R](#) for Unicode 16.0 [Reference: Section 2a of [L2/24-013](#)]

Comments: This document is an encoding proposal for an extended Egyptian Hieroglyphs repertoire, with 3,995 new characters, including 3,473 ‘Core list’ characters. (“Core” characters should be supported in widely available fonts.)

This revised proposal now includes the one additional character, located at U+143E8, which Egyptologists requested be separately encoded (see page 2 of [L2/23-254](#)). The Nov 2023 UTC

had requested a new document with a revised repertoire (including the new character) be submitted; this document fulfills that request.

2b Egyptian Hieroglyphs Database

Document:

[L2/24-035](#) Proposed Draft UAX #57, Unicode Egyptian Hieroglyph Database (Unikemet) - Suignard (= [PRI 493](#))

Recommendation: We recommend the UTC make the following disposition:

1. **Action Item** for Rick McGowan, UTC: Close [PRI #493](#) Proposed Draft UAX # 57 Unicode Egyptian Hieroglyph Database (Unikemet). [Reference: Section 2b of [L2/24-013](#)]
2. **Consensus:** Authorize a PRI for Draft UAX #57 Unicode Egyptian Hieroglyph Database (Unikemet) for 16.0. [Reference: Section 2b of [L2/24-013](#)]
3. **Action Item** for Michel Suignard, SAH: Prepare a Draft UAX #57 Unicode Egyptian Hieroglyph Database (Unikemet). [Reference: Section 2b of [L2/24-013](#)]
4. **Action Item** for Rick McGowan: Post a PRI for Draft UAX #57 Unicode Egyptian Hieroglyph Database (Unikemet) for 16.0. [Reference: Section 2b of [L2/24-013](#)]

Comments:

This Proposed Draft UAX has received copyedit feedback by the Editorial Committee. We recommend [PRI #493](#) be closed, and Michel Suignard prepare a Draft UAX (with an accompanying PRI to be opened, after the Draft is ready).

Michel Suignard reports he is fine with not zipping the Unikemet data file, so this change should be reflected in the next version of the UAX.

II. MIDDLE EASTERN SCRIPTS

3. Arabic

3a Eight More Honorifics

Document: [L2/24-002](#) Proposal to encode eight more Arabic honorifics – Pournader et al.

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Provisionally assign 8 code points U+FD90, U+FDC8..U+FDCE for Arabic Honorifics characters as described in [L2/24-002](#) [Reference: Section 3a of [L2/24-013](#)]
2. **Action Item** for Ken Whistler, UTC: Add U+FD90, U+FDC8..U+FDCE to the Pipeline (Reference: [L2/24-002](#) and Section 3a of [L2/24-013](#))
3. **Action Item** for Roozbeh Pournader, SAH: Submit a font for the eight Arabic Honorifics U+FD90, U+FDC8..U+FDCE to Michel Suignard. (Reference: [L2/24-002](#) and Section 3a of [L2/24-013](#))

Comments: This proposal comprises honorific ligatures found in fonts that the proposal author has found textual evidence for. This document proposes the encoding of eight of them. These are similar in nature to the 22 honorific ligatures already encoded in the Arabic Presentation Forms-A block. It also documents 34 other ligatures that have been found in fonts for which evidence has not yet been found.

3b Potential disunification of KASHMIRI YEH

Document: [L2/23-273R](#) Report on Kashmiri Yeh (revised) – Manish Goregaokar

Recommendation: We recommend the UTC make the following disposition:

1. **Action Item** for Manish Goregaokar, SAH: Prepare a proposal for the UTC that is a fleshed-out version of the “Stronger recommendations” section of [L2/23-273R](#). [Reference: Section 3b of [L2/24-013](#)]

Comments: Kashmiri has a letter used for palatalization, represented as U+0620 ARABIC LETTER KASHMIRI YEH. It takes on two rather distinct forms: one with a ring below ("gol"), and one with a "half" tail. The half-tail form is only found in isolated and final positions. According to the Core Spec (and many fonts), the final and isolated forms of the ringed shape is found everywhere, however real-world Kashmiri users seem to disagree with this, and as a result have come up with alternate representations for the half form at the end of words, including the Pashto letter U+06CD ARABIC LETTER YEH WITH TAIL as well as occasionally U+06C1 ARABIC LETTER HEH GOAL.

This report documents research done to understand this situation better (both by consulting written and printed material with this letter, and talking to native speakers). The main pertinent findings are that:

- The Gol form (with a ring below) is not expected or attested at the ends of words.

Whatever we decide on the overall issue, we should no longer be suggesting this.

- There is some attestation of the Taler ("half") form being found in the middle of words, in situations where it would not simply be considered a narrow space between morpheme boundaries. Attestations are found in multiple dictionaries for the same words.

The report does not yet propose a solution, but sketches out two likely solutions that can be properly proposed by a future document.

- Either we encode a separate ARABIC LETTER HALF YEH and recommend KASHMIRI YEH in final form be replaced by it
- Or, we tighten up the language in the Core Spec and mandate that KASHMIRI YEH in final and isolated form is always in its Taler ("half") form. Mid-word Taler Yeh can be achieved using a ZWNJ.

The SAH recommended the second option above (and does not currently wish to pursue disunification as proposed in the first option). If the UTC agrees, a new joining group and schematic name in ArabicShaping.txt will be needed, and Roozbeh Pournader can take an Action for that item.

III. EAST ASIAN SCRIPTS

4. CJK Punctuation

Document: [L2/23-272](#) Propose to Add Script_Extension for some CJK Punctuations --
CheonHyeong Sim

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Revise entries in ScriptExtensions.txt file to add values as documented in [L2/23-272](#), for Unicode 16.0: **Mong** to U+3001, **Mong Phag** to U+3002, and **Mong Tibt** to U+3008, U+3009, U+300A, U+300B. [Reference: Section 4 of [L2/24-013](#)]
2. **Action Item** for Roozbeh Pournader, PAG: Add **Mong** to scx of U+3001, **Mong Phag** to U+3002, and **Mong Tibt** to U+3008, U+3009, U+300A, U+300B, for Unicode 16. [Reference: Section 4 of [L2/24-013](#)]

Comments: This proposal provides evidence that several CJK punctuation marks (U+3001, U+3002, U+3008, U+3009, U+300A, U+300B) are used in Mongolian, Phags-Pa, and/or Tibetan, and requests these scripts be added to the Script_Extensions property of those characters. It also documents some other punctuation marks that are used this way but are

already `Script_Extensions=Zyyy` (Common).

IV. SOUTH AND CENTRAL ASIAN SCRIPTS

5. Kirat Rai and Tulu-Tigalari

Documents: Background proposals: [L2/22-043](#) Proposal to Encode Kirat Rai script and [L2/22-031](#) Updated proposal to encode the Tulu-Tigalari script in Unicode

Recommendation: We recommend the UTC proceed with encoding these scripts as previously proposed and accepted [Consensus: [171-C13](#) and [170-C9](#)]. We also recommend:

1. **Action Item** for Robin Leroy, PAG: Add normalization tests for all potential decompositions of U+16D68 KIRAT RAI VOWEL SIGN AI, U+16D6A KIRAT RAI VOWEL SIGN AU (proposed in [L2/22-043](#)), and U+113C5 TULU-TIGALARI VOWEL SIGN AI (proposed in [L2/22-031](#)), for Unicode 16.0. [Reference: Section 5 of [L2/24-013](#)]

Comments: Kirat Rai and Tulu-Tigalari have included some sequences that may cause problems for implementers.

Kirat Rai has:

- $E + E = AI$
- $AA + E = O$
- $O + E = AU$
 - $AA + E + E = AU$
 - $AA + AI = AU$

This means there are multiple potential routes for decomposing AU.

Furthermore, $E + E = AI$ and Tulu-Tigalari's similar $EE + EE = AI$ have decompositions into two equivalent characters, ones which are found in the endings of other canonically equivalent sequences.

None of this is forbidden by the normalization stability policies, but implementations could make assumptions that trip on this (indeed, ICU has faced some problems here).

A potential alternate route was discussed, to not encode AU and AI as atomic (or alternatively,

encode them but not include the canonical decompositions, and add DoNotUse tables). In general, we do not favor deliberately picking encodings that require DoNotUse tables, as they cause real world issues.

Dropping the atomic encoding entirely has multiple drawbacks:

- Collation order is the main drawback: AI and AU need to collate differently from their visual decompositions. [DUCET](#) does support "contractions" where a sequence can be mapped to a different collation element. However, DUCET currently only handles cases where a contraction produces a collation element equivalent to that of an existing code point (for example, for canonical equivalence), and changing this would require a major change to DUCET. This has the potential to cause problems for implementers.
- CLDR collation tailorings can support other contractions, but it is not clear if, for a minority language, such tailorings will be maintained and used by software, which would likely underserve the user community.
- Another drawback is that backspacing behavior in input methods is rarely sophisticated enough to handle things like "when backspacing two Es, backspace them as a unit" unless backed by normalization.

As such, we recommend the UTC proceed with the original encoding that contains atomic AU and AI characters with decompositions to their visual components, and that we include sufficient normalization tests to ensure implementations are appropriately handling these cases.

6. Sunuwar

Document: [L2/24-022](#) Comparison of Sunuwar font in Unicode proposal to font developed for Sunuwar Welfare Society -- Neil Patel

Recommendation: We recommend the UTC make the following disposition:

1. **Action Item** for Debbie Anderson, EDC: Add a table to the Core Spec for Sunuwar, describing the regional preferences, as documented in [L2/24-022](#). [Reference: Section 6 of [L2/24-013](#)]

Comments: Sunuwar is a new script in the pipeline for 16.0. People are in the process of developing fonts for this script, and there are some stylistic divergences between the glyph shapes used in Nepal (which is in the current proposal) and those used in Sikkim. This document highlights the divergences for U+11BCC SUNUWAR LETTER CARMi and U+11BD2 SUNUWAR LETTER SHYELE and asks that this be mentioned in the Core Spec. We do not recommend changing any of the representative code chart glyphs at this time.

7. Vedic Punctuation

Document: [L2/24-024](#) Addition of Kannada to Script Extensions of U+1CD3 VEDIC SIGN NIHSHVASA -- Srinidhi and Sridatta

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Add Kannada to the ScriptExtensions.txt entry for U+1CD3 VEDIC SIGN NIHSHVASA for Unicode 16.0 as described in [L2/24-024](#). [Reference: Section 7 of [L2/24-013](#)]
2. **Action Item** for Roozbeh Pournader, UTC: Add the value **Knda** (Kannada) in ScriptExtensions.txt for U+1CD3 VEDIC SIGN NIHSHVASA for Unicode version 16.0, based on [L2/24-024](#). [Reference: Section 7 of [L2/24-013](#)]

Comments: This document shows attestation of U+1CD3 VEDIC SIGN NIHSHVASA in Kannada. This character already has **scx** values of **Deva** and **Gran**. We agree the value should be added to ScriptExtensions.txt.

8. Mongolian

Document: [L2/24-025](#) Proposal to Encode One Manchu Letter -- CheonHyeong Sim, et al.

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Provisionally assign U+1879 as MONGOLIAN LETTER MANCHU ALTERNATE UE as described in [L2/24-025](#). [Reference: Section 8 of [L2/24-013](#)]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline for U+1879 MONGOLIAN LETTER MANCHU ALTERNATE UE as described in [L2/24-025](#) [Reference: Section 8 of [L2/24-013](#)]
3. **Action Item** for Debbie Anderson, CheonHyeong Sim, SAH: Send a font to Michel Suignard for U+1879 MONGOLIAN LETTER MANCHU ALTERNATE UE. [Reference: Section 8 of [L2/24-013](#)]

Comments: This document proposes one new character used in historical Manchu documents. The proposal author has confirmed with Kushim Jiang, one of the authors of the UTN on Mongolian, that the behavior described in the sections on shaping is in alignment with the draft UTN on Mongolian. We recommend the character for provisional code point assignment by the UTC.

V. SOUTHEAST ASIA, INDONESIA, AND OCEANIA

9. Batak

Document: [L2/23-275](#) Proposal to add an annotation to BATAK LETTER KARO BA – Ben Yang

Recommendation: We recommend the UTC make the following disposition:

1. **Action Item** for Ken Whistler, EDC: Add an annotation “also used in Toba” to U+1BC6 BATAK LETTER KARO BA. [Reference: Section 9 of [L2/24-013](#)]

Comments: When the Batak script was first encoded in Unicode, the Toba Batak language represented the syllable /ba/ with the code point U+1BC5 BATAK LETTER BA (𐌵), a letter with a slight ‘kidney bean’ shape.

In October, 2022, the First Congress on Toba Batak Culture decided to switch to using an oval-shaped character to represent /ba/. Thankfully, the Karo Batak language already uses an oval-shaped /ba/, encoded as U+1BC6 BATAK LETTER KARO BA (𐌶), which can be utilized for Toba Batak’s new form.

This proposal asks for an annotation to KARO BA, effectively recommending its use for Toba. The alternative was to recommend fonts perform this glyph change, based on the language. This was not preferred by the SAH.

VI. SYMBOLS

10. Numbers with slashes used in figured bass (musical symbols)

Document: [L2/23-277](#) Unicode request for numbers with slashes used in figured bass – Bala and Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Provisionally assign 8 digits with slash symbols used in music notation in the

range U+1D1F7..U+1D1FE, as proposed in [L2/23-277](#). [Reference: Section 10 of [L2/24-013](#)]

2. **Action Item** for Ken Whistler, UTC: Update the Pipeline with the 8 digits with slash symbols used in music notation in the range U+1D1F7..U+1D1FE [Reference: Section 10 of [L2/24-013](#)]
3. **Action Item** for Debbie Anderson and Kirk Miller, SAH: Provide Michel Suignard a font with the 8 digits with slash symbols used in music notation in the range U+1D1F7..U+1D1FE. [Reference: Section 10 of [L2/24-013](#)]

Comments: This request is for the digits with slashes that are used in figured-bass music notation in Baroque and neo-Baroque compositions. Because the placement of slashes on the digits' glyphs varies, atomic encoding is preferable to the use of U+20E5 COMBINING REVERSE SOLIDUS OVERLAY.

The names MUSICAL SYMBOL DIGIT .. WITH SLASH are used for consistency with the block, and DIGIT used for consistency with digits across Unicode (including characters which have shapes derived from digits, even if they are not digits themselves).

The Bidi class of these characters is L for consistency with other characters in the Musical Symbols block.

11. One Chess Symbol

Document: [L2/24-018](#) Unicode request for one chess symbol -- Bala and Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Provisionally assign code point U+2B74 EQUALS SIGN WITH INFINITY ABOVE as described in [L2/24-018](#). [Reference: Section 11 of [L2/24-013](#)]
2. **Action Item** for Ken Whistler, UTC: Add U+2B74 EQUALS SIGN WITH INFINITY ABOVE to the Pipeline as provisionally approved. [Reference: Section 11 of [L2/24-013](#)]
3. **Action Item** for Gavin Jared Bala, Kirk Miller, and Debbie Anderson, SAH: Submit a font with U+2B74 EQUALS SIGN WITH INFINITY ABOVE to Michel Suignard. [Reference: Section 11 of [L2/24-013](#)]

Comments: This document proposes an "equals sign with infinity above" sign, which is used in chess notation. We already have U+2BF9 EQUALS SIGN WITH INFINITY BELOW used in similar contexts. This character could theoretically be composed of an equals sign and U+1AB2 COMBINING INFINITY, but we tend to avoid composing equals signs (e.g. see \div U+2251 GEOMETRICALLY EQUAL TO).

12. Proposal to Change BidiMirroring property for U+226D

Document: [L2/23-274](#) Proposal to Change Bidi-Mirroring property for U+226D --
CheonHyeong Sim

Recommendations: We recommend the UTC make the following disposition:

1. **Consensus:** Change the Bidi_Mirrored property of U+226D NOT EQUIVALENT TO to Yes. [Reference: Section 12 of [L2/24-013](#)]
2. **Action Item** for Markus, PAG: Update the Bidi_Mirrored property of U+226D NOT EQUIVALENT TO to Yes. [Reference: Section 12 of [L2/24-013](#)]
3. **Action Item** for Robin Leroy, Manish Goregaokar, PAG: Propose changes to the Bidi algorithm, bidi properties, and the stability policy; to resolve issues found during the investigation and discussion of the document. [Reference: Section 12 of [L2/24-013](#)]

Comments: This proposal requests U+226D NOT EQUIVALENT TO (*) be marked as Bidi_Mirrored=Yes, as its use in Arabic math has the slash reversed. This would be similar to the status quo for U+2260 NOT EQUAL TO.

A tricky aspect of this is that this character decomposes to U+224D EQUIVALENT TO and U+0338 COMBINING LONG SOLIDUS OVERLAY, but neither character has Bidi_Mirrored=Yes. The stability policy [states that](#) the Bidi Mirroring properties preserve canonical equivalence. This is already a problem with U+2260 NOT EQUAL TO which has Bidi_Mirrored=Yes, but decomposes into U+003D EQUALS SIGN + U+0338 COMBINING LONG SOLIDUS OVERLAY both of which have Bidi_Mirrored=No. There are other similar characters already encoded with this behavior.

U+0338 is applied to characters other than NOT EQUIVALENT TO; in particular, it does get used with arrows like [U+219A](#). In such cases the arrow is not supposed to mirror, though as per actual usage in Arabic mathematics, the slash technically should. There is no way to represent this with the properties we currently have.

The arrows situation is similar to the situation with characters like U+2231/etc, which are integral signs with clockwise/anticlockwise marks on them. The integral sign should reverse in bidi but not the clockwise/anticlockwise marks.

Part of the issue here is that bidi mirroring lives somewhat outside of the bulk of the bidi algorithm, after the point where the structured reordered text has been handed over to a shaping engine. It is worth revisiting the wording of the stability policy and/or UAX #9 [L4](#) to fix this problem in the future.

13. Shatranj symbols

Document: [L2/24-020](#) Unicode request for shatranj symbols -- Bala and Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Provisionally assign the four Shatranj chess symbols U+1FA54..U+1FA57 as described in [L2/24-020](#). [Reference: Section 13 of [L2/24-013](#)]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to add the four provisionally assigned Shatranj chess symbols U+1FA54..U+1FA57 as documented in [L2/24-020](#). [Reference: Section 13 of [L2/24-013](#)]

Comments: We reviewed this proposal for four Shatranj chess symbols. *Alfil* ("the elephant" in Arabic) is the ancestor of the bishop in today's chess, and *ferz* is the ancestor of the queen.

The characters had earlier been proposed in [L2/16-293](#) by Garth Wallace, but the Script Ad Hoc ([L2/16-342](#)) asked that additional evidence be provided. (The proposal was later revised – but without the four proposed characters – in a joint proposal with Michael Everson, [L2/17-034R3](#), but again the evidence provided was not convincing: additional print examples were requested by the SAH in [L2/17-153](#).) In the proposal by Bala and Miller, ample evidence is provided, including contrastive evidence between bishop and alfil, as well as between queen and ferz.

Kirk Miller has already sent the font to Michel Suignard.

14. Three musical symbols

Document: [L2/23-278R](#) Unicode request for three musical symbols – Bala and Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Provisionally assign the following 3 musical symbol characters, as proposed in [L2/23-278R](#):
U+1D127 MUSICAL SYMBOL COMBINING STRESS
U+1D128 MUSICAL SYMBOL COMBINING UNSTRESS
U+1D1FF MUSICAL SYMBOL LONGA REST
[Reference: Section 14 of [L2/24-013](#)]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline with the 3 musical symbols 1D127 MUSICAL SYMBOL COMBINING STRESS, 1D128 MUSICAL SYMBOL

COMBINING UNSTRESS, and 1D1FF MUSICAL SYMBOL LONGA REST

[Reference: Section 14 of [L2/24-013](#)]

3. **Action Item** for Debbie Anderson and Kirk Miller, SAH: Provide Michel Suignard a font with the 3 musical symbols.[Reference: Section 14 of [L2/24-013](#)].
4. **Action Item** for Debbie Anderson and Kirk Miller, EDC: Propose text for Core Spec, describing behavior of glyphs for STRESS and UNSTRESS as described in [L2/23-278R](#). [Reference: Section 14 of [L2/24-013](#)]

Comments: This request is for Arnold Schoenberg's (1874–1951) stress and unstress symbols and the modern longa rest. Glyphs shown below:

- ◌ 1D127 MUSICAL SYMBOL COMBINING STRESS [SMuFL U+E4B6.] Figures 1–6.
- ◌ 1D128 MUSICAL SYMBOL COMBINING UNSTRESS [SMuFL U+E4B8.] Figures 1–6.
- ⏏ 1D1FF MUSICAL SYMBOL LONGA REST [SMuFL U+E4E1.] Figures 7–29.

The stress and unstress marks are two combining marks for musical notes akin to U+1D17C MUSICAL SYMBOL COMBINING STACCATO. They are named to match the same pattern.

Musical symbols are not unifiable with common combining marks, since their position depends on the base note shape being applied to.

The proposed glyphs for U+1D127 and U+1D128 sometimes flip when used below a note, which should probably be mentioned in the standard. There are some more shaping caveats mentioned in the proposal.

The longa rest is part of an older way of representing multimeasure rests as a combination of whole rests, breve rests, and longa rests. These other marks are already encoded.

The CCC should be 220=Below (not 230), the same as the other articulation marks.

15. Unicode request for Stein-Zimmermann quarter-tone accidentals (musical symbols)

Document: [L2/23-276](#) Unicode request for Stein-Zimmermann quarter-tone accidentals – Bala and Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Provisionally assign 12 musical symbols for Stein-Zimmermann quarter-tone accidentals in the range U+1D1EB..U+1D1F6, as proposed in [L2/23-276](#). [Reference: Section 15 of [L2/24-013](#)]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline with the 12 musical symbols for Stein-Zimmermann quarter-tone accidentals U+1D1EB..U+1D1F6 [Reference: Section 15 of [L2/24-013](#)]
3. **Action Item** for Debbie Anderson and Kirk Miller, SAH: Provide Michel Suignard a font with the 12 Stein-Zimmermann musical symbols. [Reference: Section 15 of [L2/24-013](#)]

Comments: This request is for the Stein-Zimmermann quarter-tone accidentals, the most common standard in music notation for quarter-tone accidentals, and for their extension by arrows to denote smaller alteration.

Unicode does have existing symbols for marking quarter-tones, e.g U+1D12C MUSICAL SYMBOL FLAT UP, but these do not support the most common standard for quarter-tone accidentals (which are proposed in this document).

The Bidi class of these characters should be L for consistency with the other characters in the Musical Symbols block.

VII. FEEDBACK

16. Bamum character name error

Document: [L2/24-008](#) Comments on Public Review Issues, Report ID ID20231118104342

Recommendation: PAG reviewed this and we support their recommendation (in [L2/24-009](#)) to add an alias of type “correction” for this character.

Comments: U+1680B BAMUM LETTER PHASE-A MAEMBGBIEE was encoded with a typo. The name should have been MAEMGBIEE. This feedback item asks for it to be fixed. PAG reviewed this and we support their recommendation to add an alias of type correction for this character.

17. On Provisionally Assigned Characters' Names

Document: [L2/23-233](#) Comments on Public Review Issues, ReportID: ID20231011152713

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Approve the name change for U+09FF from BENGALI LETTER ALTERNATE BA to BENGALI LETTER SANSKRIT BA. [Reference: Section 17 of [L2/24-013](#)]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline with the name change for U+09FF from BENGALI LETTER ALTERNATE BA to BENGALI LETTER SANSKRIT BA. [Reference: Section 17 of [L2/24-013](#)]
3. **Action Item** for Rick McGowan, UTC: Forward the feedback to Eduardo Marin-Silva from Section 17 of [L2/24-013](#).

Comments: The UTC provisionally assigned U+09FF BENGALI LETTER ALTERNATE BA at its January 2023 UTC meeting ([174-C20](#)), based on the proposal [L2/22-268R](#) and the SAH recommendations ([L2/23-012](#)). The character is used in an alternate orthography for Sanskrit/Pali in the Bangla script, where BENGALI LETTER BA is used to represent a "va" sound.

We agree with the feedback that the character name BENGALI LETTER ALTERNATE BA is not very clear or representative of its use. The character name should either be based on the shape (LETTER BA WITH MIDDLE DIAGONAL) or on its usage (PALI-SANSKRIT BA or SANSKRIT BA).

Due to the character's primary use in Sanskrit and the relatively low use of the Bangla script for PALI, we recommend a name change to BENGALI LETTER SANSKRIT BA, since the shape-based name would be confusable with U+09F0 BENGALI LETTER RA WITH MIDDLE DIAGONAL (𑂔).

The other two pieces of feedback concerned the names for arrows for Egyptology and the names of Sidetic characters, suggesting they should be named by their sound (if they have been deciphered):

- On the arrows for Egyptology: Changing the current descriptive names of the arrow characters does not add any value. No change is needed.
- For the Sidetic characters, names that are based on the numbering used by scholars is consistent with other scripts that are not yet fully deciphered (cf. Cypro-Minoan and Linear A). Annotations could be added to reflect the phonetic value, if later identified and deemed useful.

VIII. OTHER

18. Additions to ScriptExtensions

Document: [L2/23-280](#) Request for additions to ScriptExtensions - Pournader

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Add 51 entries to ScriptExtensions.txt as proposed in [L2/23-280](#) for Unicode 16.0. [Reference: Section 18 of [L2/24-013](#)]
2. **Action Item** for Roozbeh Pournader, UTC: Add 51 entries to ScriptExtensions.txt as proposed in [L2/23-280](#) for Unicode 16.0. [Reference: Section 18 of [L2/24-013](#)]

Comments: This proposal asks for various additions to ScriptExtensions based on research done by Roozbeh Pournader during his script exemplars project (see [L2/23-263](#)).

In that project, various characters were identified for which we can provide an exhaustive (or sometimes almost exhaustive) list of scripts with which they are used. The quality of the information provided for the proposed additions matches the quality of the information for existing characters with a Script_Extensions property. We think they don't need any further research and should be added to ScriptExtensions.txt at the earliest opportunity.

19. DoNotEmit.txt

Document: [L2/24-021](#) Do Not Emit (16.0.0) -- Pournader

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus:** Add DoNotEmit.txt to the UCD for Unicode 16.0 as described in [L2/24-021](#). [Reference: Section 19 of [L2/24-013](#)]
2. **Action Item** for Roozbeh Pournader, SAH: Add DoNotEmit.txt to the UCD as described in [L2/24-021](#). [Reference: Section 19 of [L2/24-013](#)]
3. **Action Item** for Ken Whistler, PAG: Update UAX 44 to include the new DoNotEmit.txt as described in [L2/24-021](#). [Reference: Section 19 of [L2/24-013](#)]

Comments: This proposes Unicode publish a DoNotEmit.txt file in UCD, containing characters and character sequences that should not be emitted or generated in newly authored text, alongside the preferred equivalent. The information is based on various tables and other information in the Core Spec and the NamesList not previously available in a machine readable format. It contains a classification of the different kinds of such sequences, for example [Indic_Atomic_Consonant](#), for "Sequences that look like an Indic consonant but should be avoided in representing that consonant".

Software can use this data to clean up text that it receives or generates, or to present alternatives for users.

20. Requirements and Process for Changing Script Status for Identifier Use

Document: [L2/24-019](#) Requirements and Process for Changing Script Status for Identifier Use -- Asmus Freytag

Recommendation: We've reviewed this document and agree with the Unicode Properties & Algorithms group's recommendation to start following these guidelines.

Comments: [UTS 39](#) defines [Identifier_Status](#), which in part depends on script status as defined in [UAX 31](#). Whereas UAX 31 recommends whether characters are allowed in an identifier-like grammar in the first place, UTS 39 recommends which sequences in identifiers are best promoted, and is more fluid. This gets used, e.g., to determine which scripts are actively supported in public domain names by registrars (including the root zone). So far, there have been no clear guidelines on how scripts can advance from Excluded/Limited_Use to Recommended.

This document proposes a concrete set of guidelines, primarily based on whether a script is used by a community in daily activities. We recommend PAG be the first contact for those wishing to change the status of a script, and PAG should decide where to post such guidelines.

B. DOCUMENTS NOT REQUIRING UTC ACTION

21. Old Hungarian

Document: [L2/24-003](#) Additional comments on [L2/21-115](#) (Proposal for modifying the Old Hungarian block) -- Deborah Anderson for Viktor Kovács

Related documents:

[L2/21-115](#) Proposal for modifying the Old Hungarian block -- Bence FEHÉR, József Álmos KATONA (SAH response: [L2/21-130](#))

[L2/21-246](#) Proposal for a compromise of the recent Old Hungarian proposal - Marin-Silva (SAH response: [L2/22-023](#))

[L2/22-285](#) Proposal for a compromise of the recent Old Hungarian proposal -- Viktor Kovács

[L2/23-067](#) Article on Old Hungarian letters for Q, X, Y -- Viktor Kovács

Recommendation: This is just an FYI to the UTC.

Comments: This document raises topics from [L2/21-115](#), providing reasons for not including '500', the modern letters Q, W, X, Y, DZ, and DZS, and a few other topics. The only actionable request is found at the end of the document: it requests an "alias" (=annotation) to the two NIKOLSBURG OE characters (U+10C9D and U+10CDD) be changed from "also used for ü" to "also used for ü by Ferenc Súlyom from the end of the 20th century." However, the SAH did not think it necessary to include this annotation, which captures fine detail on the history of the character. Michael Everson, co-author of the successful Old Hungarian proposal, agreed the annotation is not needed.

22. Cuneiform

Document: <https://www.unicode.org/reports/tr56/tr56-2.html>

Comments: This is just an FYI. SAH looked at the updates and had no further comment.

23. Ethereum symbol

Document: [L2/24-023](#) Proposal to add the Ethereum symbol to Unicode -- Josh Rudolf, Justin Drake

Recommendation: We recommend the UTC take no action.

Comments: The SAH reviewed this proposal to add one currency symbol for the cryptocurrency Ethereum. We did not find clear evidence of a distinct symbol used in public text interchange to warrant encoding a new character at this time. In particular, we did not see a clear distinction from the Ethereum logo.

The distinction between a logo and a non-logo symbol isn't always clear. The proposal states, "The Ethereum symbol is a simplified version of the Ethereum logo." But in reviewing the [Ethereum brand assets](#), we see that the symbol is, in fact, one of the brand-identifying variants of the logo. It is evident that the only reason the authors have for wanting the proposed character to be encoded, rather than using some other symbol such as U+27E0, is so that the symbol used to denote the Ethereum currency is the Ethereum *logo*. By long-standing policy Unicode does not encode logos. For this reason, the SAH recommendation to the Unicode Technical Committee is against encoding the proposed symbol.

We recommend that the Ethereum community use U+27E0 LOZENGE DIVIDED BY HORIZONTAL RULE or a similar character, as the evidence provided shows has already been done in practice.

24. Bété

Document: [L2/23-279](#) Bété (Ivorian alphabet) working document – Adam Yeo et al.

Recommendation: We recommend the UTC take no action.

Comments: This is a FYI to the UTC. It is a working document on and is an update to [L2/19-044](#) Bété Script Working Documents.

C. IN PROCESS

[L2/23-216](#) Revised request to change glyphs of four Kannada characters -- Srinidhi A, Sridatta A