

TO: UTC
FROM: Deborah Anderson, Ken Whistler, Roozbeh Pournader, Lisa Moore, Liang Hai, Chris Chapman, and Richard Cook
SUBJECT: Recommendations to UTC #155 April-May 2018 on Script Proposals
DATE: 28 April 2018

The Script Ad Hoc group met on 9 February 2018, 16 March 2018, and 23 April 2018 in order to review proposals. The following represents feedback on proposals that were posted in the Unicode document registry at the time the group met.

Editorial

1. Feedback for Editorial Committee

Document: [L2/18-025](#) General feedback for the editorial committee – Eduardo Marín Silva

We reviewed this document, which contained several sets of comments directed to the Editorial Committee. Below are the comments.

CJK Symbols and Punctuation

- #1-#17 (except #12 and #14): The names list editor will take the suggestions under advisement.
- #12: Discouraging the use of a character is not an editorial decision, but involves UTC approval. A separate document would need to be submitted, and generally a strong reason should be provided (beyond “it serves no purpose but to be confusable”).
- #14: Formal name aliases must be approved by the UTC. In our opinion, formal name aliases here are not warranted.

General Punctuation

- #1, #2, and #4 is not information that should be included in the names list. If the author can identify text in the Core Spec that is incorrect or under-specified, he may suggest wording for consideration by the Editorial Committee.
- #3 is a cross-reference that the editor will take into consideration.

Supplemental Punctuation

- #1: In our opinion, the cross reference back to ASCII parentheses is not necessary.
- #2-#5: The cross-references suggested will be taken into consideration by the names list editor.

Miscellaneous Technical

- The proposed names list suggestions will be reviewed by the names list editor, and some may be incorporated.

Combining Diacritical Marks, Combining Diacritical Marks for Symbols, and Letterlike Symbols

- The names list editor will take the suggestions under advisement; some may be incorporated.

CJK Compatibility Ideographs

- #1: Providing a list of the CJK Unified Ideographs in the CJK Compatibility Ideographs block with a (proposed) formal alias is very unlikely to be approved by the UTC or WG2. Such a change would have implications for the Name property and require a significant change in the tooling.

The information on the CJK Unified Ideographs for the Compatibility Ideographs is already tracked by means of characters properties and by entries in the UniHan database.

C0 Controls and Basic Latin and C1 Controls

- The suggested additions will be remanded to names list editor. It was noted that the names list is not intended to be encyclopedic and including all similar-looking characters as cross-references is not necessarily helpful to the reader.

Number Forms

- The suggestions to the Number Forms names list will be turned over to names list editor to review. He will incorporate those that are deemed reasonable.

Mongolian

- The suggested changes for Mongolian (i.e., to remove all the contextual forms) will not be adopted. The entire model for Mongolian is still under discussion, and information on contextual forms is still needed for users and implementers.

Recommendations: We recommend the UTC remand the applicable suggestions above to the names list editor for consideration.

2. Public Review Comments

Document: [L2/18-009](#) Comments on Public Review Issues (October 13, 2017 - January 20, 2018)
The following topics were not taken up during the UTC or have not already been addressed.

Feedback on Encoding Proposals

a. Miao - Marín Silva

The comment from Eduardo Marín Silva was to change the name of MIAO SIGN CONSONANT MODIFIER BAR to MIAO SIGN COMBINING CONTRAST OF ARTICULATION. In the original proposal for Miao additions, L2/17-345, the name for the character was MIAO SIGN NUKTA. After discussion at the October 2017 UTC it was changed to MIAO SIGN CONSONANT MODIFIER BAR, a name that was agreeable to the proposal author and UTC members.

We reviewed the comments from Adrian Cheuk in [L2/18-069](#). We agree with the Adrian Cheuk: a functional name is not needed, nor is “COMBINING” needed. In our assessment a change in the name is not worth revisiting.

Recommendations: We recommend the UTC note the feedback but take no action.

Error Reports

b. Changes to Indic_Syllabic_Category and Indic_Positional_Category – Corbett COMBINING ASTERISK ABOVE

We reviewed this request and recommend Roozbeh Pournader review whether U+20F0 COMBINING ASTERISK ABOVE should have Indic_Syllabic_Category=Cantillation_Mark and Indic_Positional_Category=Top. The submitter noted that COMBINING ASTERISK ABOVE has scx={Deva Gran Latn} because it is used as a svara marker.

Recommendations: We recommend the UTC remand this topic to Roozbeh Pournader to review and make any changes if warranted.

c. Underspecified Soyombo vowel signs (and Zanabazar Square) – Corbett
Underspecified Zanabazar Square vowel signs - West

David Corbett requested the Core Spec explicitly specify the syllabic structure of Soyombo and Zanabazar Square. Current Unicode editorial staff resources are limited, so if someone pulls together all the necessary information in a fully specified description of the Soyombo/Zanabazar square syllabic structure, the UTC should review the text and then remand it to the Editorial Committee for inclusion in the Core Spec.

Andrew West requests the Core Spec spell out the encoding order of the Zanabazar Square script vowel signs and vowel length mark to aid implementers and users. In our view, such information would be welcome, but a document with such information should be submitted to the UTC for approval. After approval by the UTC, text can be added to the Core Spec.

Recommendations: We recommend the UTC relay to the feedback authors that additional documentation needs to be submitted before text can be added to the Core Spec. The authors are welcome to submit such documents.

Note: This feedback has been forwarded to the author of the Soyombo and Zanabazar Square proposals, Anshuman Pandey.

d. Typo in the section on Kayah Li – Corbett

The feedback notes an apparent error on page 653 in 16.9 Kayah Li of TUS, where the vowels are written <σ'> and <u'> instead of <σ> and <u>.

Recommendations: We recommend Julie Allen make the changes, if feasible.

Note: This change has been incorporated in Unicode 11.0.

e. Indic_Syllabic_Category of U+0C80 – Corbett

The comment from Corbett is that the Indic_Syllabic_Category for U+0C80 KANNADA SIGN SPACING CANDRABINDU should be Bindu, based on [L2/14-153](#).

Recommendations: We recommend U+0C80 KANNADA SIGN SPACING CANDRABINDU should have the Indic_Syllabic_Category Bindu, after getting verification from Roozbeh Pournader that “Bindu” is correct.

f. Script_Extensions of U+1CF2 VEDIC SIGN ARDHAVISARGA (add Tirhuta) – Corbett

Based on [L2/11-175R](#), which noted use of U+1CF2 VEDIC SIGN ARDHAVISARGA in Tirhuta, Corbett requested Tirhuta be added to the scripts in ScriptExtensions for U+1CF2 VEDIC SIGN ARDHAVISARGA.

Recommendations: We recommend Tirhuta be added to the set of scripts in the ScriptExtensions property for U+1CF2 VEDIC SIGN ARDHAVISARGA.

g. Unspecified order of Meetei Mayek vowel signs – Corbett

The feedback noted that the order of vowel signs in the “Abbreviations” section of Chapter 13.7 (p. 541) is not specified. In addition, the final sentence in the same section is technically incorrect. In our assessment, the text requires technical review so more accurate wording can be proposed.

Recommendations: We recommend the UTC respond to thank the author of the feedback and suggest he (or others) consider reviewing relevant materials, check with experts, and provide new wording.

h. Indic_Syllabic_Category of U+A8B4 SAURASHTRA CONSONANT SIGN HAARU – Corbett

Comments received noted that U+A8B4 SAURASHTRA CONSONANT SIGN HAARU is a modifier letter that can be followed by a vowel sign or virama. It is currently assigned the Indic_Syllabic_Category of Consonant_Final but should perhaps be changed to Consonant_Medial. We agree with the comment.

Recommendations: We recommend the Indic_Syllabic_Category of U+A8B4 SAURASHTRA CONSONANT SIGN HAARU be changed from Consonant_Final to Consonant_Medial.

i. Syloti Nagri dvisvara and anusvara – Corbett

The feedback mentions that U+A802 SYLOTI NAGRI SIGN DVISVARA is underspecified in the Core Spec. Corbett also pointed out the position of U+A80B SYLOTI NAGRI SIGN ANUSVARA and U+A802 SYLOTI NAGRI SIGN DVISVARA is unusual for an Indic script, and their use should be documented in the standard. The comments also noted that *dvisvara* should have the Indic_Syllabic_Category assignment “Vowel_Dependent” and Indic_Positional_Category as “Top.”

In our view, this feedback requires technical review before any full recommendations can be made.

Recommendations: We recommend the UTC create an Action Item for Peter Constable and the Editorial Committee to review the comments, so the feedback is not lost. In addition, we recommend Roozbeh Pournader should be given an Action Item to review the Indic_Syllabic_Category Vowel_Dependent and Indic_Positional_Category designations, and make changes if warranted.

3. Core Spec Changes

Document: [L2/17-423](#) Changes to Core Specification for Indic scripts – Srinidhi and Sridatta

The recommendations below are a follow-up to the extensive Script Ad Hoc comments in [L2/18-039](#) Recommendations to UTC #154 January 2018 on Script Proposals (pp. 4-6). The recommended actions in this section (Core Spec Changes) are actions for the Editorial Committee. Greyed-out items have already been taken care of and are incorporated in Unicode 11.0.

§1 Sharada

The request and ad hoc recommendations were to change the glyphs in §15.3 for *jihvamuliya* + KA and *jihvamuliya* + KHA.

Recommendations: We recommend Julie Allen be notified of this change, Rick McGowan give Liang Hai the font, and Liang Hai provide the corrected glyphs to Julie. The shapes should be:



(*jihvamuliya* with KA)



(*jihvamuliya* with KHA)

Note: *This change has been incorporated in Unicode 11.0.*

Note: There is an Action Item for Liang Hai and Roozbeh Pournader to review all the Indic syllabic categories for the existing *jihvamuliya* and *upadhmaniya* characters (Action Item 154-A21).

§2 Kannada

The request and ad hoc recommendations were change the glyph for Kannada *hka* (in §12.8 of TUS).

Recommendations: As above for the Sharada glyph corrections, we recommend Julie Allen get the corrected glyph from Liang Hai (via Rick McGowan). The correct shape should be:



Note: This change has been incorporated in Unicode 11.0.

§3 Sora Sompeng

The original request and ad hoc recommendations were to change the description of Sora Sompeng from “abugida” to “alphabet” in §15.14, and Table 6-1.

Recommendations: We recommend Ken Whistler work with Julie Allen to make this correction.

Note: This change has been incorporated in Unicode 11.0.

§4 VEDIC SIGN NIHSHVASA

The authors requested the description of U+1CD3 VEDIC SIGN NIHSHVASA in §12.1 of the Core Specification be aligned with the code chart annotation, which specifies that U+1CD3 “separates sections between which a pause is disallowed.” The current description in §12.1 says the character “indicates where a breath may be taken.” The authors recommend the wording “Separates sections of Sama Vedic singing between which a pause is disallowed.” The ad hoc agreed that wording change was needed.

Recommendations: We recommend Lisa Moore work with Julie Allen on improved wording for section §12.1 of the Core Spec (Vedic Extensions, p 467).

Note: Text has been incorporated in Unicode 11.0 addressing this issue.

§5 Ligature Forms for Ra + Vocalic Liquids

The authors recommend the Core Specification be revised to mention that the sequence /r vocalic_r/ can appear graphically in two forms in scripts besides Devanagari:

Recommendations: We recommend Liang Hai add a short paragraph after Table 12-4, indicating the pattern of behavior for *ra* and *vocalic r* is observed in other scripts, such as Kannada, etc.

Note: Text has been incorporated in Unicode 11.0 addressing this issue.

§6 Brahmi

The authors requested a glyph correction for figure 14-1 to:



They also recommend a minor correction in the text (with one fix of a typo in their suggested wording). The ad hoc agreed with these changes.

The corrected text should read:

U+11002 BRAHMI SIGN VISARGA is used to write syllable-final voiceless /h/. The velar and labial allophones of /h/; that is, /χ/ and /ϕ/, followed by voiceless velar and ...

Recommendations: We recommend Julie Allen ask Andrew Glass for the glyph update for figure 14-1, insert the corrected glyph, and make the text change as noted above.

Note: This change has been incorporated in Unicode 11.0.

4. Script Extensions

Document: [L2/17-424](#) Changes to ScriptExtensions.txt for Indic characters for Unicode 11.0 – Srinidhi and Sridatta

The recommendations below were in the Script Ad Hoc comments in [L2/18-039](#) (p. 7), but were not discussed at the January 2018 UTC. The Script Ad Hoc has discussed §1 further at the February 2018 Script Ad Hoc meeting. (The questions posed by the Script Ad Hoc on §2.3 has been forwarded to the authors. The remaining items under §2, §3 and §4 from [L2/17-424](#) have already been addressed.)

§1 Devanagari




DEVANAGARI SIGN PUSHPIKA

Comments: The authors request Newa and Kannada be added to the set of scripts in the ScriptExtensions property for U+A8F8 DEVANAGARI SIGN PUSHPIKA, and when Nandinagari and Tigalari are encoded, similar adjustments be made the set of scripts in the ScriptExtensions property for U+A8F8 to account for usage by those scripts.

It was noted by the Script Ad Hoc that the character U+A8F8 DEVANAGARI SIGN PUSHPIKA does not currently have a ScriptExtensions property. The shape of the *pushpika* represents a flower, but it does not interact with neighboring characters. The *pushpika* might be compared to OM or the dandas, whose shapes tend to get harmonized on a per script basis.



What is the general approach going forward? Should a *pushpika* be encoded on a per-script basis, or should one *pushpika* to be used across the scripts (using fonts to display the preferred shape)? Or should the approach be based on the grouping of closely related scripts (i.e., one *pushpika* for the northern scripts and one for the southern scripts)?

Compare:

Northern scripts: Devanagari  Newa  Nandinagari  (a southern development of the northern Nagari style)

Southern scripts: Kannada  Tigalari  ¹

¹ Note the different shape in Tigalari proposal L2/17-378, where it is called OM ALANKAARA:

Proposed shape  ; image from figure 23: 

Recommendations: We recommend the UTC discuss the general approach to *pushpika*, and based on the outcome of the discussion, make any relevant changes to the ScriptExtensions property.

AFRICA

5. Medefaidrin

Document: [L2/18-142](#) Medefaidrin corrections for 11.0 - Moses Ekpenyong, et al.

Comments: We reviewed this document, which relayed feedback to the UTC on the Unicode 11.0 beta code charts and names list for Medefaidrin. The document provided evidence for the requests.

The following summarizes the input and our views:

- The feedback identified true errors in four character names: the case pair for U+16E56/U+16E76 MEDEFAIDRIN LETTER HP (which should have been MEDEFAIDRIN LETTER H), and U+16E57/U+16E77 upper and lowercase MEDEFAIDRIN LETTER NY (which should have been MEDEFAIDRIN LETTER NG).
- In our view, the need to modify the open *o* character's name, U+ 16E5A/U+16E7A MEDEFAIDRIN LETTER OE, is debatable, since the open *o* does not have a consistent name in the names list.
- The request to change the name for U+ 16E99 MEDEFAIDRIN SYMBOL AIVA to MEDEFAIDRIN SYMBOL OR is not, in our view, required, as an annotation can be used.
- Two suggested modifications to glyphs (U+16E56 MEDEFAIDRIN CAPITAL LETTER HP [sic] and U+16E6B MEDEFAIDRIN SMALL LETTER I) can be made at a later time.

Recommendations: We recommend the UTC decide how to handle the corrections to the names for the following four characters (> show corrected names):

16E56 MEDEFAIDRIN CAPITAL LETTER HP > H
16E76 MEDEFAIDRIN SMALL LETTER HP > H
16E57 MEDEFAIDRIN CAPITAL LETTER NY > NG
16E77 MEDEFAIDRIN SMALL LETTER NY > NG

Correcting the names in the UCD and other files risks a possible delay in the release of 11.0. Alternatively, the name corrections could be handled as Formal Name Aliases of the type "correction."

The UTC may decide to modify the name for U+ 16E5A/U+16E7A MEDEFAIDRIN LETTER OE, but we don't feel it is necessary. We recommend the UTC remand to the names list editor a request for an annotation to U+ 16E99, and to Deborah Anderson to follow up on glyph changes to U+16E56 and U+16E6B.

SOUTHEAST ASIAN

6. Thai (Thai Noi)

Document: [L2/18-072](#) Towards a comprehensive proposal for Thai Noi / Lao Buhan script - Ben Mitchell

Comments: We reviewed this document, which provided very useful information about the Thai Noi script. The document explicitly states that Thai Noi is not an extension to the Thai alphabet (page 2). The document includes suggestions and information that the author of [L2/18-041](#) Request to Add Thai Characters (WG2 N4927) could usefully incorporate when preparing a proposal for Thai Noi.

Recommendations: We recommend the UTC quickly review this document, but request it be forwarded to the author of [L2/18-041](#) (Nitaya Kanchanawan), WG2, and ISO/IEC JTC1/TC46/WG3.

Note: The Script Ad Hoc reviewed the documents listed below on Thai (Thai Noi) and submitted its recommendations in [L2/18-070](#) Feedback on ISO/TC 46/WG3 N2631 ISO/CD 20674-1 "Information and documentation - Transliteration of scripts in use in Thailand - Part 1: Transliteration of Akson-Thai-Noi."

[L2/18-041](#) Request to Add Thai Characters (WG2 N4927) - Nitaya Kanchanawan

[L2/18-042](#) Information and documentation - Transliteration of scripts in use in Thailand - Part 1: Transliteration of Akson-Thai-Noi (WG2 N4927A) - TC46 / WG3

[L2/18-043](#) Results on ISO CD 20674-1: Information and documentation - Transliteration of scripts in use in Thailand - Part 1: Transliteration of Akson-Thai-Noi (WG2 N4927B) - TC46 Secretariat

Related documents:

[L2/18-068](#) Thai-Noi Transliteration (WG2 N4939) – Martin Hosken

SOUTH AND CENTRAL ASIA

7. Vedic

Document: [L2/18-075](#) Proposed Property Changes for Ardhavisargas - Srinidhi A, Sridatta A

Comments: We reviewed this document which was a response to the request by the Script Ad Hoc recommendations in [L2/18-039](#). The Script Ad Hoc recommended the authors prepare a separate document exploring the interaction of U+1CF2 VEDIC SIGN ARDHAVISARGA with characters before and after it, and to address the question whether its general category should be changed to a non-combining spacing character. (Sharma had argued for gc=Lo for visarga, anusvara, and related characters including ardhavisarga in L2/09-343.) Does VEDIC SIGN ARDHAVISARGA participate in re-ordering or shaping?


The document [L2/18-075](#) includes various new examples, beyond those that were in [L2/17-424](#), and also examines U+1CF3 VEDIC SIGN ROTATED ARDHAVISARGA. Based on the examples, the two ARDHAVISARGAs do not appear to interact with surrounding characters or cause re-ordering.



We agree with the authors that the general category (gc) should be changed from Mc to Lo, and Line Break from CM to AL. With the change in the gc, the representative glyph should have the dotted circle removed.



Recommendations: We recommend the UTC change the general category of U+1CF2 VEDIC SIGN ARDHAVISARGA and U+1CF3 VEDIC SIGN ROTATED ARDHAVISARGA from Mc to Lo, and Line Break from CM to AL, and modify the glyphs to remove the dotted circle. We further recommend Roozbeh Pournader and Liang Hai determine the best IndicSyllabicCategory for U+1CF2 and U+1CF3, and check on the additions of Bengali, Kannada, Telugu and Tirhuta to ScriptExtensions.txt.



8. Bengali

Document: [L2/18-035](#) Encoding model issues with the Vedic *gomukha* characters – Sharma






Comments: We reviewed this document, which requests guidance from the UTC on the encoding model for *gomukha* characters in Bengali, raised specifically by the Bengali form , cited in §4.1.4 of [L2/17-](#)

[098](#). The author noted that the Bengali form  is a variant of  U+A8F3 DEVANAGARI SIGN CANDRABINDU VIRAMA.



The form appears to be the sequence U+1CEA  VEDIC SIGN ANUSVARA BAHIRGOMUKHA and U+1CED  VEDIC SIGN TIRYAK, but which character should be used to represent the dot?


The Vedic Extensions section of Chapter 12 of the Core Spec (p. 468) mentions that the *gomukha* characters may be combined with  DEVANAGARI SIGN CANDRABINDU (U+0901) or  DEVANAGARI SIGN ANUSVARA (U+0902). With evidence now in Bengali for *gomukha* characters, which characters should be used for the combining marks? Should the Devanagari characters be used?

The author provides different options:

1. Use U+0982 BENGALI SIGN ANUSVARA  and U+0981  BENGALI SIGN CANDRABINDU, noting that BENGALI SIGN ANUSVARA is not a “dot above” shape.
2. Use U+0902  DEVANAGARI SIGN ANUSVARA and U+0901  DEVANAGARI SIGN CANDRABINDU. Sharma notes that Bengali has its own *candrabindu* at U+0981: 
3. Encode a separate “dot above” character for Bengali.
4. Use U+0307 COMBINING DOT ABOVE and U+0310 COMBINING CANDRABINDU in the Combining Diacritical Marks block
5. Encode two new characters in the Vedic Extensions block, a script=inherited “dot above” character and a CANDRABINDU.
6. Encode only a new “dot above” character with script=inherited in Vedic Extensions, since *candrabindu* is uniform across the Indic scripts, but the glyph for the *anusvara* may vary.

The following comments were raised during discussion:

- Is the “dot above” only found in Devanagari?
- Some members of the Script Ad Hoc felt option #2 or possibly #6 might work, but there was hesitation in adding more dots (options 3, 5, 6) without careful consideration of the viable alternatives.
- Why couldn't U+09FC  BENGALI LETTER VEDIC ANUSVARA be used, on the model of  U+A8F3 DEVANAGARI SIGN CANDRABINDU VIRAMA?

Recommendations: We recommend the UTC review this document, seeking input from those members with Indic rendering engines and the Bengali user community. We also suggest the author work with Srinidhi and Sridatta, authors of [L2/17-098](#), and consider why U+09FC  BENGALI LETTER VEDIC ANUSVARA would not be suitable.

9. Malayalam

Document: [L2/18-015](#) Proposal to encode the END OF TEXT MARK for Malayalam – Srinidhi and Sridatta

Feedback:

[L2/18-034](#) Feedback on L2/18-015 proposing Malayalam punctuation mark – Sharma

[L2/18-009](#) Comments on Public Review Issues from Eduardo Marín Silva

[L2/18-145](#) Feedback on name for MALAYALAM END OF TEXT MARK (L2/18-015) - Anderson

Comments: We reviewed the proposal, which requested one character, MALAYALAM END OF TEXT MARK.

The feedback from Sharma ([L2/18-034](#)) is in support of the character. He provides transliteration of the examples in the proposal, which clarifies the name “end of text,” but noting that the mark predominantly appears alongside dandas. Sharma prefers the name MALAYALAM SIGN PUSHPIKA (*pushpa*=flower), since its usage is very similar to U+A8F8 DEVANAGARI SIGN PUSHPIKA. The Devanagari character also appears alongside dandas.

In our view, the name MALAYALAM SIGN PUSHPIKA is acceptable, and preferable to END OF TEXT MARK (or MALAYALAM SECTION MARK as suggested by Eduardo Marín Silva). The authors of the proposal are not against the name PUSHPIKA as noted in L2/18-145, although they note that the shape technically does not resemble a “flower.”

Recommendations: We recommend the UTC accept the character U+0D53 MALAYALAM SIGN PUSHPIKA, after first discussing the name. However, we recommend discussion about the overall approach to the *pushpika* character in Indic scripts be addressed first (discussed above, #4. Script Extensions, §1 *Devanagari*).

10. Sinhala

Documents:

[L2/18-060](#) Proposal to encode the CANDRABINDU for Sinhala – Srinidhi and Sridatta

[L2/18-079](#) Feedback on Sinhala *candrabindu* ([L2/18-060](#)). – Anderson

Comments: We reviewed the proposal L2/18-060, which requested one character, SINHALA SIGN CANDRABINDU. At our initial review in February, the Script Ad Hoc noted that the examples all come from one publication, and wondered if the *candrabindu* it appears in other publications? Other questions posed by the Script Ad Hoc included: Is the *candrabindu* in common usage, or only rare? If it is rare, why couldn’t experts use U+0310 COMBINING CANDRABINDU? The February meeting ended with a recommendation that this proposal be circulated for input from other experts and the Sri Lanka representatives. Deborah Anderson followed up with Sri Lanka representatives and posted the response in L2/18-079.

At the March meeting of the Script Ad Hoc, we reviewed the short response document L2/18-079, which provided feedback from Sri Lanka experts on the proposal to encode a Sinhala *candrabindu* ([L2/18-060](#)). The comments attest to the use of the *candrabindu* in Sanskrit materials from at least the 19th century, citing more than one source.

The use of U+0310 COMBINING CANDRABINDU from the Combining Diacritical Marks block was not recommended by a member of the Script Ad Hoc, since its use should be limited to Latin, Cyrillic, and Greek scripts.

Based on corroboration by experts in Sri Lanka, we agree there is valid reason to encode SINHALA SIGN CANDRABINDU. An annotation in the names list or a comment in the Core Spec citing the character’s use in Sanskrit materials might be worth considering.

Recommendations: We recommend the UTC accept U+0D81 SINHALA SIGN CANDRABINDU.

11. Pallava

Document: [L2/18-083](#) Preliminary proposal to encode Pallava – Pandey

Comments: We reviewed this preliminary proposal. In our view, the current direction of the proposal is reasonable, and we encourage the author to continue his work on it.

Two minor notes:

- The code points in Character Data in section 3 (page 5) starts at “1E030”, as in the current Roadmap, but the chart and names list start at “1E300.”
- The caption on page 11 should be from “NA...LLLA” instead of “DHA...LLLA”

Recommendations: We recommend the UTC members review this document at their leisure and forward any comments to the proposal author.

Note: This feedback has already been forwarded to the proposal author.

12. Sirmauri

Document: [L2/18-085](#) Preliminary proposal to encode Sirmauri – Pandey

Comments: We reviewed this preliminary proposal. The following are comments that arose during discussion:

- Provide a chart comparing Sirmauri with Takri
- Propose a different name, so this script won’t be confused with the living language of the same name
- Remove the virama, unless specimens can be provided in support of a virama

Recommendations: We recommend the UTC members review this document at their leisure and forward any comments to the proposal author.

Note: This feedback has already been forwarded to the proposal author.

13. Khwarezmian

Document: [L2/18-010](#) Proposal to encode the Khwarezmian script in Unicode – Pandey

[Note: The proposal seen by the Script Ad Hoc has been revised based on comments below. The script has also been renamed “Chorasmian” and given a new document number [L2/18-164](#). The Script Ad Hoc did not review [L2/18-164](#).]

Comments: We reviewed a revised version of the proposal for Khwarezmian, and have the following comments:

- Add a new character for what is now called “final aleph” (cf. Nabataean FINAL ALEPH)
- If there is a real need for differentiating WAW and YOD, please provide a detailed analysis of their differences and in which kind of source material they occur, and address the potential encoding issues for cases when there is no visual distinction in the source material. Also differentiate their glyphs for such a case.
- Create two tables in place of one table in 4.1 (pp. 4-5): a dual-joining table (cf. Table §9.8 in *TUS*) and a right-joining table (cf. Table §9.9 in *TUS*). For the right-joining table, remove the initial and medial forms.
- Note in the proposal whether the script has ligatures.
- Discuss the distinction between a *letter* that is in isolated/initial/medial/final position in terms of cursive joining vs. a letter in those positions in a *word* (i.e., a letter can be in the middle or final position of a word but may not be cursive-joining medial or final; instead, it could be a cursive-joining isolate, for example).

Recommendations: We recommend the UTC review this document and forward any comments to the proposal author.

Note: This feedback has already been forwarded to the proposal author.

14. Khitan Small Script

Document: [L2/18-121](#) Cluster Formation Model for Khitan Small Script – West et al.

Comments: We reviewed this document, which proposes a different model for handling the Khitan Small Script (KSS) clusters than had been proposed earlier.

Controversy surrounding the model has delayed progress of the script in the ISO balloting process. (The repertoire of the script in the latest ballot, PDAM 2.2, did not include any format control characters, as they had been removed in an earlier ballot.)

According to the proposal author, Khitan Small Script has a fixed set of cluster shapes – following either the “Cluster A” or “Cluster B” layout:

<i>Cluster A</i>	<i>Cluster B</i>
AB	A
CD	BC
E	DE

The following summarizes the background on the KSS model discussion:

- The model put forward in the final KSS proposal ([L2/16-245](#)) involved use of one (of two) format characters prefixed to a cluster: one to indicate the block be rendered with the “Cluster A” layout (above), and a second one to indicate rendering in the alternative structure, “Cluster B.”
- The WG2 meeting in September 2016 recommended in [L2/16-338](#) that two format characters—a horizontal joiner and a vertical joiner—be used between characters in a cluster, as has been proposed for Egyptian hieroglyphs.
- In response, Andrew West in [L2/16-296](#), felt that inserting vertical and horizontal format characters in every cluster would impose an unnecessary burden on end users, and basically restated the model originally proposed (a format character be placed before the first character of the structure with a double initial clusters, and a second format character appear before the first character when it starts with a single initial cluster).
- The current document recommends no format control characters, instead inserts Combining Grapheme Joiner (CGJ) between the first and second characters. The cluster sequence is terminated by EOL or any non-KSS graphic character or CGJ. Clusters in the text stream are separated by a space character, and space characters separate single standalone characters in the text stream.

The following were comments raised during discussion:

- The new proposal no longer proposes a prefixed format character, but now recommends a model with an infixed character. We agree with the author that an infixed model is the best approach.
- CGJ has specific functions in Unicode (*TUS* §23.2): when inserted into a sequence of combining marks, it blocks the canonical re-ordering of the combining marks (such as in Hebrew, see *TUS* §9.1, and in Arabic, see [UTR 53 Unicode Arabic Mark Rendering](#)). CGJ may also be used to affect

the collation of adjacent characters, but its usage as such is discouraged in UTS #10. However, because the proposed use of CGJ in KSS falls outside of prescribed usage, it may not work as intended in software and could cause problems in rendering engines.

- Concerning the “unnecessary burden” for users caused by having to type horizontal and vertical joiners between characters: An IME can be created that won’t require users to type a joiner between each character.

Recommendations: We recommend the UTC review this document and forward any comments to the proposal author, including those above.

15. Mongolian

Document: [L2/18-107](#) Summary of suggestions and proposals for improvements to the Mongolian phonetic model – Roozbeh Pournader

Comments: We reviewed this summary of suggested improvements to the current Mongolian phonetic model which arose out of the Mongolian meeting in San Jose, CA, from April 3-5, 2018.

The different groups of attendees were invited to identify their three top priorities, which were drawn from their presentations. The priorities are listed below, with our comments. Detailed background on the priorities will be relayed during the UTC. Other topics (besides the top three priorities) could be taken up later.

From Badral Sanlig (top three priorities based on: [L2/18-101](#) Analysis of the graphetic model and improvements to the current model):

Priority 1. The feminine forms of MONGOLIAN LETTER QA and MONGOLIAN LETTER GA either should be separately encoded as separate characters, or as fixed variants with FVS1.

Comments: A proposal is needed with rationale and evidence. It was noted that the grapheme for the feminine forms of QA and GA is already encoded as an Ali Gali character (U+1889 MONGOLIAN LETTER ALI GALI KA), so the Ali Gali character could be one alternative.

Priority 2. We should fix the NNBS character as soon as possible or encode a new suffix connector for Mongolian.

Comments: NNBS currently has line-break property GL[UE], indicating it prohibits line breaks before and after. However, Mongolian allows a line-break before NNBS in limited contexts. In addition, UAX #14 and sections §6.2 and §13.5 of *TUS* state NNBS is typically displayed with one third the width of a normal space character, but there is evidence NNBS is stretchable (depending up on typographical style) and is not always less than a normal space. Also, it was noted that NNBS loses its width (i.e., it disappears) over a line-break.

We recommend the following sections of *TUS* be reviewed and revised by the Editorial Committee regarding the behavior of NNBS:

- UAX #14 Unicode Line Breaking Algorithm
- §13.5 Mongolian of the Core Spec (p. 534 of version 10.0)
- §6.2 of the Core Spec

The UTC may wish to reconsider GL as the LB property for NNBSB.

Priority 3. Stylistic and historical characters and variation sequences in the Mongolian block need to be cleaned up and potentially reorganized into the Ali Gali section of the Mongolian block.

Comments: We recommend no action at this time.

From Menghejiya (top three priorities based on: [L2/18-103](#) Revised Proposal for minimizing the current phonetic code)

Priority 1. All positional variants of a letter, including its default form, should have a corresponding variations sequence. (Presently, the default form of a letter in a position does not have a variations sequence specified for it.)

Comments: Adding standardized variations sequences for default shapes should be constrained to only those cases where it is required to make a distinction. If a proposal is submitted listing the needed positional variants, it can be reviewed by Unicode experts. We recommend Liang Hai work with Menghejiya and others in preparing such a document.

Priority 1.1 (A corollary to Priority 1) A fourth Free Variation Sequence needs to be added for a dotless medial form of GA and potentially other positional variants for other letters, including encoding a new Free Variation Selector 4.

Comments: The addition of a fourth FVS seems generally inadvisable. Use of existing VS, such as U+FE00 VARIATION SELECTOR-1, should be employed, unless a clear case can be made why using VS instead of a fourth FVS is problematic. We recommend Liang Hai work with Menghejiya and others on this.

Priority 2. There are various problems with the NNBSB. We should either:

- A. Add a new Mongolian Suffix Connector character in the Mongolian block; or**
- B. Specify that NNBSB does not affect Mongolian shaping, and its only role is to make sure there is no word break at that position.**

Comments: We recommend no action at this time.

Priority 3. We need a control character for: 1. For limiting the effect of feminine or masculine vowels in a word when the word has two or more roots of different genders. 2. For breaking automatic ligatures

Comments: For (1), existing VS should be employed. For (2), ZWJ could be used. Additional documentation may be needed in the Core Spec to address the specific issues cited in this item.

From Enkhdalai Baatar (top three priorities based on: [L2/18-105](#) Mongolian Script in Unicode)

Priority 1. A new character should be added for the feminine form of QA and GA, disunified from the currently encoded QA and GA. This would help reduce the number of rules needed for shaping Mongolian.

Comments: See comments above for Badral Sanlig.

Priority 2. Stylistic and historical variants should be implemented using fonts, not variations sequences. Unicode encoding should not include style variants, they should be removed so the rules in Unicode would be simplified and shortened.

Comments: In our view, including stylistic and historical forms is a lower priority than other changes required to get modern Mongolian to work. We recommend the UTC continue to work with the various parties.

Priority 3. MVS and the three Free Variation Selectors should be replaced by just one Free Variation Selector (with potentially two or more of it used in a row).

Comments: The Script Ad Hoc sees this as an issue to be resolved by the design of the input mechanism, rather than encoding.

Additional information based on discussion with experts (pp. 3-4):

1. U+1806 MONGOLIAN TODO SOFT HYPHEN is not limited to Todo, but is also used in Hudum to separate compound words.

Comments: We recommend the UTC remand this item to the Editorial Committee, so the information about the usage of U+1806 MONGOLIAN TODO SOFT HYPHEN is captured either in the names list or the Core Spec.

2. U+202F NARROW NO-BREAK SPACE (also known as suffix connector)

Comments: It was noted that NNBS can be preceded by non-Mongolian text. In such cases, contextual behavior should still work.

Recommendations: We recommend the UTC discuss the document from the April 2018 Mongolian Working Group Meeting 2. We recommend the UTC remand the NNBS topic to the Editorial Committee for changes to *TUS*.

16. Old Uighur

Document: [L2/18-126](#) Preliminary proposal to encode Old Uighur – Pandey

Comments: We reviewed this preliminary proposal for the Old Uighur script.

The following comments were raised during discussion:

- We recommend the spelling “Uyghur” be used throughout the proposal. Include a note at the beginning of the document about the spelling.
- §5.4.1
 - The medial and initial forms of GIMEL and HETH look identical. Investigate whether they do indeed share the same form. If so, the proposal should recommend that for medial

GIMEL/HETH, users should use the medial form of (either GIMEL or HETH), and include similar text for the initial forms, explaining this is because the characters share the same shape.

- Investigate if the medial and initial forms of BETH and YOD have the same shape. We recommend YOD be made right-joining, and for medial and initial forms of YOD (if it can be confirmed BETH/YOD have identical glyphs in these positions), users should employ BETH. As a general rule-of-thumb, if two distinct forms (“a” and “b”) from an earlier era later merged into one shape, identify in the proposal which character to use for which era.
- For the final forms: provide more information so a decision on whether separate encoding is warranted.
- §5.3
 - Decide on the canonical forms for the charts. Note: The current code charts for Mongolian reflect how the script is implemented in fonts, that is, in a horizontal direction, with glyphs rotated.
 - Verify if there is a strong tendency to typeset the script in a particular direction, drawing on existing documents.
 - Discuss how the text appears in a Latin script context.
 - Identify changes in script orientation through time.
 - Keep the developers of CSS informed on this script.
- §5.4.2 Combining signs
 - Provide an explanation on the names with “DOT” (since the representative glyph is a stroke), mentioning other scripts (such as Arabic)
 - Base the names on the canonical orientation of the glyphs in the charts (which could affect whether the names with “BELOW” and “ABOVE” are appropriate)
 - The combining signs act as nuktas, i.e., as diacritics used to create new letters. Indicate this feature in the discussion, as it affects the combining class properties. Make a case why the combining marks should not be treated as atomic units with joining groups, rather than as nuktas. Why are they treated differently from Arabic or Manichaean? Since Old Uyghur is an historic script and it isn’t likely that new letters will be found, consider whether encoding a DOTTED GIMEL, etc., with a joining group would be simpler.
 - Follow up on the diacritic signs mentioned in Erdal and Clark.
- §6.1 Provide more detail on the shift of the baseline after MEM, with examples showing the glyph interaction
- Provide fuller information on figures 8-19 (i.e., describe what the text contains, date, etc.)
- Reword or remove figure 39 (page 56), “Comparison of Uighur letters with Unicode Mongolian letters.” The comparison between Old Uyghur graphemes with the currently encoded Mongolian characters (which are not directly related to graphemes) is confusing and misleading.
- Consider including a reference to the Uyghur-Mongolian draft standard in China.

Recommendations: We recommend the UTC review this document and send comments to the author.

Note: This feedback has already been forwarded to the proposal author.

17. Tibetan

Document: [L2/18-078](#) Deprecation of 3 Tibetan Characters - Élie Roux

Comments: We reviewed this request to deprecate three Tibetan characters: U+0F00 TIBETAN SYLLABLE OM and two head marks (U+0F02 TIBETAN MARK GTER YIG MGO -UM RNAM BCAD MA and U+0F03 TIBETAN MARK GTER YIG MGO -UM GTER TSHEG MA).

The Unicode Standard (D13 in section 3.4) states that deprecated characters typically are those characters that pose significant architectural problems or cause implementation problems. In our view, the three characters do not rise to this level.

The characters all appeared in Unicode 2.0 in 1996, and hence were encoded before normalization was first introduced (Unicode 3.1). All the OM characters in Unicode are “Lo”, and none are decomposed. Like the Mongolian *birgas*, the two headmarks are encoded as atomic symbols (gc=“So”), which is why collation doesn’t decompose them.

The author is invited to propose a short annotation for the names list on the usage of these characters or text for the Core Spec that describes their use or provides guidance on the representation of the characters.

Recommendations: We recommend the UTC discuss this proposal and send the feedback above to the author, with any other comments.

18. Zanabazar Square

Document: [L2/18-132](#) Proposal to encode two additional Zanabazar Square letters (WG2 N4945) – Andrew West

Comments: We reviewed this proposal for two Zanabazar Square cluster-initial letters LA and SA, which correspond to Tibetan head letters LA and SA that appear in conjuncts. The two proposed characters correspond to already encoded character U+11A3A ZANABAZAR SQUARE CLUSTER-INITIAL LETTER RA.

The author identifies a distinction between Sanskrit transcription and Tibetan transcriptions in Zanabazar Square, where the Sanskrit examples show the two consonant glyphs with a gap between the glyphs, and the Tibetan transcription depicts a compressed ligature, as shown in the chart on page 1:

	RA + KA	LA + KA	SA + KA
Sanskrit Syllable			
Tibetan Syllable			

The following comments arose during discussion:

- Clarification is needed on how to handle cases that fall outside the “typical” examples shown on page 1 (above). How should the following be encoded: Figure 3 (compressed and not ligated), Figure 5 (halfway compressed and not all clearly ligated), Figure 7 (halfway compressed and ligated), and Figure 11 (compressed and not ligated). If no clear guidelines are provided, encoding these two new characters might result in encoding LA and SA in two ways, with no way to distinguish them.

- Provide examples showing contrastive use in a single source.

It was noted that the proposal modified the glyph for the currently encoded U+11A3A CLUSTER-INITIAL LETTER RA, changing the dotted box to a dotted circle. The dotted box, indicating special rendering, appears in the Soyombo code chart for the four cluster-initial letters. We believe the dotted circle in place of the dotted box is an oversight of the author.

Recommendations: We recommend the UTC review this document and send comments to the author, including those above.

MIDDLE EAST

Arabic

19. Arabic for Hausa

Document: [L2/18-094](#) Proposal to encode additional Arabic script characters for Hausa - Lorna Evans, Andy Warren-Rothlin

Comments: We reviewed this proposal for two new Arabic script characters needed to represent Hausa. The proposal contains solid evidence in support of the characters. The proposal also includes helpful information for the Core Spec on the characters needed for the Hausa orthography and how the glyphs in that orthography should appear.


A correction is needed for the name of U+08C3 in the table on page 2 and elsewhere in the proposal. We recommend the name include “ABOVE”, hence ARABIC LETTER GHAIN WITH THREE DOTS ABOVE. (In the properties and on page 3, the character’s name is ARABIC LETTER AIN WITH DOT ABOVE AND THREE DOTS ABOVE, which also needs to be corrected.)



Recommendations: We recommend the UTC approve the following two characters:
 U+08C3 ARABIC LETTER GHAIN WITH THREE DOTS ABOVE
 U+08C4 ARABIC LETTER AFRICAN QAF WITH THREE DOTS ABOVE

20. Luri




Document: [L2/18-061](#) Proposal to include Luri alphabets - Mohammad Mogoei, Lateef Shaikh

Comments: We reviewed this proposal for five characters for the Luri language, used in Western Iran. Two characters can already be represented in Unicode:

For  ARABIC LETTER ALEF WITH SUKUN ABOVE
 use the sequence <U+0627 ARABIC LETTER ALEF, U+0652 ARABIC SUKUN>

For  ARABIC LETTER FARSI YEH WITH SMALL INVERTED V ABOVE
 use U+063D  ARABIC LETTER FARSI YEH WITH INVERTED V

The remaining three characters, listed below, need evidence showing them in newspapers, magazines, or other publications which demonstrate widespread usage.

	U+08C4	ARABIC LETTER ALEF WITH SMALL INVERTED V ABOVE
	U+08C5	ARABIC LETTER KEHEH WITH SMALL INVERTED V ABOVE
	U+08C6	ARABIC LETTER LAAM WITH SMALL INVERTED V ABOVE

Recommendations: We recommend the UTC discuss this document and relay feedback to the authors.

21. Hindko

Document: [L2/18-032](#) Proposal to include Hindko alphabets - Lateef Sagar Shaikh

Comments: We reviewed this proposal, which requested five Arabic characters used to write the Hindko language in northern Pakistan. The characters appear to be well-justified.

The following comments were raised during discussion:

- If the characters need to be kept together as a set, the proposed code points, U+08BE..U+08C2, are acceptable. Another alternative is to put the first character in the hole at U+08B5, and the other four at U+08BE..U+08C1.
- The letter names should be modified as follows in order to be consistent with other Arabic character names:
 - ARABIC LETTER PEH WITH SMALL V
 - ARABIC LETTER TEH WITH SMALL V
 - ARABIC LETTER TTEH WITH SMALL V
 - ARABIC LETTER TCHEH WITH SMALL V
 - ARABIC LETTER KEHEH WITH SMALL V
- We recommend the glyph for KEHEH WITH SMALL V should lose the “s” between the “v” and the letter’s body (left, below), so it appears like the glyph on page 2 (right, below):



page 1



page 2 (recommended glyph by Script Ad Hoc)

The glyph on page 1 (found also on page 3) is a valid alternative, but the glyph with the “s” could confuse users of the standard about the character’s identity.

- The proposal should include ArabicShaping.txt data. However, we do not think the proposal should be blocked if it does not include this information.

Recommendations: We recommend the UTC accept the five characters, after deciding on the code point locations, and request the author update the proposal with the code points, names, and the glyph change (as noted above).

EAST ASIA

CJK

22. IDCs

Document: [L2/18-012](#) Proposal of 4 IDCs - Yang, et al.

Feedback: [L2/18-048](#) Regards to L2/18-012 – Taichi Kawabata

Comments: We reviewed the proposal, which requested four IDCs, and feedback from Kawabata in L2/-048. The input from Kawabata notes the ambiguity of unary operators and deems the proposed IDCs to be of limited benefit. We agree with Kawabata’s assessment.

The following were other points raised during discussion:

- “Using [the 4 new IDCs] may reduce the quantity of unnecessary and irregular components for IDS” (p. 1)

Discuss the benefit of the proposed additions, argue in favor of them, and demonstrate the cost-benefit for including them. Is the benefit being proposed sufficient to justify moving beyond compatibility that was originally intended for the IDCs?

- Are there other operators the authors will be proposing?
- The first IDC, 2FFC IDEOGRAPHIC DESCRIPTION CHARACTER SURROUND FROM RIGHT, presents no syntactic problems. If it is useful, it would be acceptable, in our opinion.

The second IDC, 2FFD IDEOGRAPHIC IDEOGRAPHIC DESCRIPTION CHARACTER SURROUND FROM LOWER RIGHT, is problematic, with incorrect syntax. In this case, we do not feel there is a need for this IDC. Couldn't 2FFC be used?

2FFE and 2FFF fundamentally change the syntax, and hence are very problematic. They would require constraints in the syntax, since they could introduce several ways to represent the same structure.

- How would the following be represented with the current set of IDCs, and how is what is proposed better?

U+6C37	氷	☐氷、
--------	---	-----

Recommendations: We recommend the UTC review these documents and forward comments (including those above) to the proposal authors.

23. UAX #45 Additions

Document: [L2/18-064](#) Proposal to add 2 new UAX#45 characters – Chan

Comments: We reviewed this proposal, which requested the addition of two characters to UAX #45. Evidence is provided, and we agree adding the two characters to UAX #45 is advisable.

Recommendations: We recommend the UTC add the two characters to UAX #45 as requested, after discussing this document.

24. Extension B

Document: [L2/18-063](#) Proposal to remove the UCS2003 representative glyphs from the Extension B code charts – Lunde

Comments: We reviewed this document, which requested the removal of the UCS2003 representative glyphs from Extension B code charts and provides five reasons in support of this request. Since the change would involve a structural change to the charts, affecting a change in tooling, we defer this proposal to the UTC and the Editor of ISO/IEC 10646, Michel Suignard.

Recommendations: We recommend the UTC discuss this proposal.

25. CJK Punctuation

Document: [L2/18-073](#) Proposal to add standardized variation sequences from various punctuation – Lunde (Supersedes: [L2/18-013](#) Proposal to add standardized variation sequences for digits and various punctuation – Lunde)

Comments: We reviewed this document and its predecessor, [L2/18-013](#), which was the second part of [L2/17-056](#) Proposal to add standardized variation sequences. It requests 63 standardized variation sequences for 43 characters. In our view, this proposal is outside the purview of the Script Ad Hoc.

Recommendations: We recommend this proposal be discussed by the UTC.

26. Shuishu

Documents: [L2/18-131](#) Comments on Shuishu in PDAM2.2 text – Suzuki

[L2/18-133](#) Additional Comments on Shuishu in PDAM2.2 text (WG2 N4946) – Suzuki

Comments: We reviewed these documents. The Script Ad Hoc supports the reasoning given by Suzuki. We are of the opinion that the Shuishu proposal is not yet mature, and hence encoding at this time is not advisable.

Recommendations: We recommend the UTC review these documents.

NUMBERS

27. Rejang Numbers

Document: [L2/18-081](#) Preliminary proposal to encode Rejang Numbers – Pandey

Comments: We reviewed this preliminary proposal. In order to be able to properly review this document and discuss the model, actual examples of the signs in running text are needed. Are there actual users of Rejang numbers?

Recommendations: We recommend the UTC members review this document at their leisure and forward any comments to the proposal author.

Note: This feedback has already been forwarded to the proposal author.

SYMBOLS

28. Postal Mark

Document: [L2/18-058](#) Proposal to encode the POSTAL MARK ENCLOSED IN DOWN POINTING TRIANGLE - Eduardo Marín Silva

Comments: We reviewed this proposal, which requested one character, POSTAL MARK ENCLOSED IN DOWN POINTING TRIANGLE, that appears in the Morisawa and Sha Ken glyph sets. This character was originally proposed as part of a set in [L2/18-004](#).

The proposal provides clear justification and evidence of the character as a distinct symbol. We suggest the code point be in the Enclosed Alphanumeric Supplement block at U+1F1AD, so it can be near other circled or boxed characters that serve as generic symbols.

We recommend the proposal include a clear example of the proposed glyph in the body of the text (besides the example from a tweet) and cite the font which contains the glyph in the proposal summary form.

Recommendations: We recommend the UTC accept the character, after discussing the code point and name, but recommend the author make modifications to the proposal as suggested above.

29. DPRK compatibility symbols and punctuation (KPS 9566)

Document: [L2/18-057](#) Request to provide explicit rationales for exclusion of symbols in the KPS 9566 standard - Eduardo Marín Silva

Comments: This document was a response to UTC #154 feedback on the author’s earlier proposal ([L2/18-004](#) Proposal to reconsider compatibility symbols and punctuation used in the DPRK).

Recommendations: We recommend the UTC note this feedback.

30. Moon Symbols

Document: [L2/18-062](#) Revised proposal for dealing with the moon symbols problem – Karlsson

Comments: We reviewed this proposal, which takes up the question of how to represent the moon symbols (U+263D..U+263E and U+1F311..U+1F319) as viewed from the Southern Hemisphere vs. the Northern Hemisphere. The problem was raised in [L2/17-304](#), a document that proposed Variation Sequences. As noted in the October 2017 Script Ad Hoc recommendations in [L2/17-384](#), use of VSes to represent the Southern vs. Northern Hemisphere views of the moon was not an appropriate use of VSes, in our opinion.

In [L2/18-062](#), the author proposes extensive aliases and cross-references be added to the names list. In our opinion, the names list is not the appropriate place for such detail. The 11.0 chart has added wording before U+1F311, below the subheading “Moon, sun, and star symbols” to address the topic:

Use of the moon symbols is typically reversed in the Southern Hemisphere, with the waxing shapes used for the waning part of the cycle and vice versa. The character names are not intended for astronomical precision, but simply describe a particular shape displayed.

The author is invited to send feedback on the 11.0 wording (above) as part of the 11.0 beta.

Recommendations: We recommend the UTC forward the above comments, as well as any other comments to the proposal author.



SCRIPT AD HOC RECOMMENDATIONS CARRIED OVER FROM JANUARY 2018
 ([L2/18-039](#) Recommendations to UTC #154 January 2018 on Script Proposals)

SOUTH ASIA

31. Malayalam Vedic Anusvara

Document: [L2/17-276](#) Proposal to encode MALAYALAM LETTER VEDIC ANUSVARA– Srinidhi and Sridatta

Feedback document: [L2/17-419](#) Feedback on Malayalam Vedic Anusvara – Shriramana Sharma

Comments: We reviewed [L2/17-276](#), the proposal document for MALAYALAM LETTER VEDIC ANUSVARA, with evidence of its use.

The evidence is solid for this character. We noted that the term “VEDIC” is contained in two other characters: U+09FC BENGALI LETTER VEDIC ANUSVARA and U+1135E GRANTHA LETTER VEDIC ANUSVARA, so there is precedence for the proposed name.

We also reviewed the feedback document by Sharma, which provides rationale to move MALAYALAM LETTER VEDIC ANUSVARA from U+0D50 to U+0D04. We agreed with Sharma’s rationale for the move. The new location (U+0D04) would fill a hole near the beginning of the Malayalam block, and place the character near other similar phonetic characters. As noted by Sharma, it also leaves U+0D50 open in the event an OM character is found. In addition, it may be important to maintain a correspondence with other Indic scripts that typically have OM in that position.

Recommendations: We recommend the UTC approve MALAYALAM LETTER VEDIC ANUSVARA, with the glyph as given in [L2/17-276](#), but decide on the code point location.

32. Syloti Nagri

Document: [L2/17-418](#) Encoding model to represent conjuncts in Syloti Nagri – Srinidhi and Sridatta

Background document: [L2/05-130](#) Encoding Model for Syloti Nagri Conjoining Behaviour - Constable

Comments: We reviewed this document, which reconsiders the model for rare cross-cluster ligatures and false conjuncts in Syloti Nagri. The authors recommend the UTC (1) discuss the plain-text representation of cross-cluster ligatures and false conjuncts and make changes if necessary to §15.1 Syloti Nagri in the Core Spec, and (2) consider a change in the Indic Syllabic category for U+A806 SYLOTI NAGRI SIGN HASANTA, from Pure_Killer to Virama.

Of the various cases cited in Table 1 and 2, of particular interest are the following:

Table 1 (with Burmese model, with virama)

Case	Example sequence	Conjoining display with specialty font	Bad non-conjoining display with basic font	Correct non-conjoining display
“false” conjunct” C + C	< 𑌔 , virama, 𑌔 > (“kot”)	𑌔𑌔	𑌔𑌔	𑌔𑌔
“false conjunct” with spacing vowel mark	< 𑌔 , 𑌔 , virama, 𑌔 , 𑌔 > (“bibir” – same sequence as “bibri”)	𑌔𑌔𑌔	𑌔𑌔𑌔	𑌔𑌔𑌔

Table 2 (with ZWJ model)

Case	Example sequence	Conjoining display with specialty font	Conjoining display with basic font
“false” conjunct” C + C	< क्त , ZWJ, क्त > (“kot”)	क़्त	क़्त
“false conjunct” with spacing vowel mark	< क्, ी, क्, ZWJ, ी, क् > (“bibir” – distinct sequence from “bibri”)	कीक़ी	कीक़ी

The following points were raised during discussion:

- The UTC had earlier agreed with Peter Constable’s analysis in [L2/05-130](#), i.e., the model for the script is one with a virama, with no conjoining behavior (like Burmese).
- For special forms, such as rare cross-cluster ligatures or false conjuncts, [L2/05-130](#) recommended OpenType features or ZWJ. The only way to be able display such special conjoining behavior today would be with joiners.
- As noted in the document, the cross-cluster ligatures or false conjuncts are rare, and occur in handwritten documents, but not in modern printed sources. The authors feel use of ZWJ should not be used in these contexts, but they should only be handled at the font level with OT features.
- Are all the cited conjoining forms orthographically significant?
- Two approaches to the problem could be to either handle this behavior in higher-level protocol or add a new conjunct-forming letter.
- Are conjuncts more common than ligatures? If the script ligates often, then a stacker may not be necessary.

In our view, the present Indic syllabic category, Pure_Killer, is acceptable. If there is a need for representing either conjuncts or false conjuncts (i.e., intrasyllabic ligatures) in plain text, it should be demonstrated. If such a case is justified, then a different solution should be considered, such as a new stacker character.

Recommendations: We recommend the UTC review this proposal, and discuss it with Peter Constable, the author of L2/05-130.

33. Khojki

Document: [L2/17-307](#) Proposal to encode two characters in Khojki – Srinidhi and Sridatta

Comments: We reviewed this document, which requested two Khojki characters. The following were comments raised during the discussion:

- Is the evidence based on just one source?
- Provide a translation of the text in figure 2.
- How is the independent vowel represented?

Recommendations: We recommend the UTC review this document and send comments, including those above, to the author.

34. Tigalari

Document: [L2/17-378](#) Preliminary proposal to encode Tigalari script -- Vaishnavi Murthy K Y, Vinodh Rajan

Background documents:

[L2/17-182](#) Comments on encoding the Tigalari script – Srinidhi and Sridatta

Related documents:

[L2/17-411](#) Letter in support of preliminary proposal to encode Tigalari - Guru Prasad

[L2/17-422](#) Letter to Vaishnavi Murthy in support of Tigalari encoding proposal - A. V. Nagasampige

[Note: Information has been received that a competing proposal will be submitted in the future.]




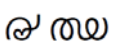
Comments: We reviewed this proposal, which addresses some of the points raised by Srinidhi and Sridatta in L2/17-182 and the Script Ad Hoc comments in L2/16-342. The proposal now includes properties.

The following comments were made in a review of this proposal:

- In order to provide a track record of Tigalari proposals, we recommend the authors cite earlier versions of this proposal at the top of the document (on page 1), i.e., “[This document] replaces L2/16-241.” It would also be useful to cite L2/17-182 and the ad hoc comments L2/16-342 within the proposal (such as in §4), so comments from other documents can also be tracked.
- Currently Tigalari is only allocated 6 columns on the [Roadmap](#), but the proposal has 8. The current proposal appears to be based on the Malayalam block, which was itself based on ISCII. A better use of space would be to move the digits left one column, put the letters for vocalic rr and ll and vowels signs for vocalic l and ll, in the usual Sanskrit order (or, for any Dravidian-specific letters, in the appropriate Dravidian relative order).
- §5.1. Include a chart in the body of the proposal showing the representation of independent vowels (which appear to have distinct graphic pieces), similar to the one for Malayalam (in place of a reference in footnote 14):

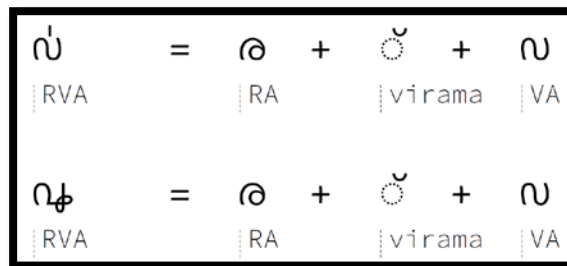
For	Use	Do Not Use
ഇ	0D08	<0D07, 0D57>
ഉ	0D0A	<0D09, 0D57>
ഐ	0D10	<0D0E, 0D46>
ഓ	0D13	<0D12, 0D3E>
ഔ	0D14	<0D12, 0D57>

- §5.1 (pages 6-7) LETTER E and LETTER O:
 - Do research to clarify whether the proposed *e/o* are used currently, and appear in educational materials. If they are in use, then they are eligible for encoding. If not currently being used, then they can be added later.
 - Create a table to show how *e* and *o* appear in different documents in figures 42, 43, and 44, and provide examples of their use in texts by those authors.
- §5.2 (page 8): In the section directly under “For the TIGALARI AI LENGTH MARK...U+113C8 should be used” the text should read “EE-vs + EE-vs” (not E-vs)
Similarly in §9 on page 29, on the top, E-vs should be corrected to EE-vs.

- §5.2 (page 8): The TIGALARI AU LENGTH MARK is an acceptable character to encode; it is a useful graphetic unit for the writing system. However, it should be moved to a position at the end of the range of other dependent vowels
- §5.2 (page 9): The proposal states that the vowel signs U and UU change shape, depending upon the consonant to which they combine. Provide a comprehensive list of the consonant and vowel combinations that attach graphically, noting any exceptions. Provide this information for all the dependent vowels. (Examples appear to be provided on p. 7 of L2/17-182)
- §5.2 (page 9): For the alternate form of the vowel sign U , describe when it occurs. Does it appear consistently in the same manuscript? Is the form interchangeable with other forms of U? This information can help identify whether it is best handled on the font-level (as a font style, for example), or whether it needs to be distinguished in plain text, and hence separately encoded.
- §5.2 (page 10): Under “2. Placed below and right ligating”, the text reads “While...LA + Virama combinations can be used here, it is recommended not to as the characters are not canonically equal.” The wording “not canonically equal” here is unclear; can the authors explain? L2/17-182 states the combination is a conjunct. We recommend the authors incorporate the information from L2/17-182, page 9 (sections 2 and 3)
- §5.4 Virama (page 12ff): For many historical scripts, separate viramas for a killer and a stacker may be advisable, in order to avoid over-reliance on joiners and non-joiners, which can be very confusing for users. It should be noted that using Malayalam as a model is not necessarily advisable, since Malayalam is a different orthography. We recommend two viramas be proposed: a vowel-silencing virama (“killer”, with a visible mark) and a stacker (with no visible mark).
- §5.4 (pages 14-15): The ligature forms of *k*, *t*, *tt*, and *n* resemble Malayalam *chillus*, though the model is simpler than for Malayalam. We recommend separately encoding 4 *chillu* forms.
- §5.5 (page 16, top): Remove the paragraph “The Malayalam model..” with the two lines of examples. If the authors feel strongly that the alternate forms should be represented, demonstrate that the distinction is important, and provide use cases. (Note: This topic could be proposed separately at a later point, with the authors’ suggestions on how to represent the forms in text [i.e., by a stacking virama and a joiner or non-joiner].)
- §5.5 (page 18) 4. Ligating Special Characters: We would recommend a Tigalari *dot reph* be separately encoded (see below, §8.5).
- §5.5 (page 19, top): Represent  as *<dot reph, ka, killer>*.
- §6.1 (p. 20): Alternate Glyph Shapes. The third form of Tigalari Letter II, , is separately encoded in Malayalam (U+0D5F MALAYALAM LETTER ARCHAIC II), based on L2/12-225. Consider whether it should be separately encoded (or not).
- §6.1 (p. 21): The shapes of the consonants and alternates in the list on page 21 seem to vary considerably, e.g., JHA . Could they different be characters? Provide the rationale why the two JHA forms are not proposed as two characters (though pronounced similarly). Provide better discussion for other consonants which have vastly different shapes. Incorporate information from L2/17-182 (pp. 2-4), as applicable.
- §7 Digits (p.22): Is the numbering system as represented in figure 22 (and the chart on p. 22) rare, or is it used widely in historical documents? According to the authors of L2/17-182, the manuscripts in figures 20 and 21 seem to be in the Malayalam script. Is the numbering system

as proposed used in schools today? If not, we recommend the authors hold off proposing them until more information on them can be provided.

- §8 (p. 23) *Candra Anunaasaika*:
 - The name *Candra Anunaasaika* varies from the more usual Indic pattern of naming. We recommend it be spelled *anunasaika* (cf. U+0901 with the informative alias *anunasaika*). Is the character ever referred to as *candrabindu*?
 - In order to determine whether the mark is spacing or is more like TELUGU SIGN CANDRABINDU (a right-side mark), demonstrate whether the Tigalari sign functions as part of the syllable or is a spacing mark between syllables.
- §8.2 (p. 24): Add the names and code points to the “Double danda” reference (i.e., U+0965 DEVANAGARI DOUBLE DANDA), if no script-specific dandas are proposed.
- §8.2 (p. 24-25): We suggest the authors put the three characters, OM ALANKAARA, SHRII ALANKAARKA (*shrīi* symbol), and PUSHPA ALANKAARA into a chart. In accompanying text, provide justification for PUSHPA ALANKAARA (as the text currently mentions U+2055 – or is that only meant to be a tip to font creators?). The two characters OM ALANKAARA and SHRII ALANKAARKA (*shrīi* symbol) seem to be good candidates for encoding.
- §8.3 (p. 25 and figure 25): *Tiddu* mark (Correction mark) is not justified, and we recommend the authors remove this section. The mark is described as an editorial convention used to insert missing text, or to mark imperfections in the text. There are no guidelines as to the mark’s placement, and the *tiddu* was often added later to texts. If a mark is required, the DEVANAGARI CARET at U+A8FA could be employed, or markup could be used. A *de novo* creation of a standardized editorial convention is not the realm of plain text, in our view.
- §8.4 (p. 26) Vedic tone marks:
 - Although Vedic tone marks are not proposed here, we encourage the authors to consider whether existing Vedic marks could be used, since they may appear in related manuscripts in the same area of India.
 - Mark the tone marks in the figure 26.
- §8.5 (p. 26) *Reph*: In order to represent the vertical tick at the beginning of a syllable, we recommend a *dot reph* be separately encoded, which avoids use of joiners/non-joiners. A *stacker* character should be used to represent the post-consonant or sub-based form. Hence, the top example would use a *dot reph*, but a *stacker* character would be used to represent the bottom example:



- §8.7 (p. 28): If script-specific *dandas* are not proposed, add a firm statement that users should employ Devanagari *dandas*.
- §9 (p. 29): Fix the typos for E-vs in the second line (needs a dotted circle), and in the fifth line, correct EE-vs to E-vs.
- §10.1 (p. 30) Collation: Remove *tiddu*, add 4 *chillus*, *dot reph*, *stacker* and *killer* characters. Note: LLA appears twice (correct one to LLLA).

- §10.2 (p. 30ff): OM ALANKAARA should be given the General_Category value Lo (as in Devanagari). (As above, remove *tiddu*, add 4 *chillus*, *dot reph*, *stacker* and *killer* characters. Incorporate other suggested moves as noted.)

Recommendations: We recommend the UTC forward the above comments to the authors. We also suggest the authors work with the authors of L2/17-182 (Srinidhi and Sridatta).

Note: This feedback has already been forwarded to the proposal authors.

35. Amaragannada scripts

Document: [L2/17-186](#) Introducing the Amaragannada scripts – Srinidhi and Sridatta

Comments: We reviewed this document, which introduces a set of approximately 10 historic scripts for possible encoding. The scripts, used primarily in Northern Karnataka in SW India, were devised to protect the religious literature of Kodekal Basavanna from destruction by Islamic rulers. The major language used was Kannada.

A review of this document raised the following questions:

- How widespread are these scripts?
- Are the scripts named?
- If there is a one-to-one mapping between these scripts and Kannada (such as in figures 5, 6, 7 and 8), why should these be encoded? Why wouldn't a font-change suffice? Provide a convincing use-case. It was noted that while a cipher *could* become a full-fledged scripts (i.e., Thaana), many do not. Separately encoding a cipher for Kannada may end up obscuring the true content of the Kannada text.

Recommendations: We recommend the UTC members review this proposal, and send comments to the proposal authors.

Garay

36a. Proposal

Document: [L2/16-069](#) Proposal for encoding the Garay script in the SMP – Everson

Comments: We reviewed this proposal for Garay. The following summarizes points raised in the discussion:

- Based upon the evidence presented, such as in figure 7 (below), the VOWEL SIGN E and above-comma diacritical mark on MBA, NGGA, and NDA are graphically the same. As a result, we recommend the characters MBA, NGGA, and NDA be removed, and instead they be represented as a base letter and VOWEL SIGN E.

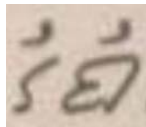



Figure 7:

- Based on the evidence, the dot above should be separately encoded. Besides NJA and NGA, the examples show other letters with a dot above.

Figure 24: \ddot{h} \ddot{s} and \dot{h} \dot{s} (in caption) and in the image, above the left leg of GA: 

Hence, we recommend the removal of NJA and NGA, and instead they be represented as a base letter and combining dot above.

- In §2.2.3 Gemination, the text reads:

“The gemination mark (◌̇) is written above a consonant letter and above the VOWEL SIGN E where that applies to a letter (◌̇). In encoded text, for searching and matching, perhaps the best practice would be to encode the gemination mark (bound to the consonant) first and the VOWEL SIGN E afterward, even though the presentation form appears to be the opposite.”

The wording in the second sentence should be revised, as the vowel mark should be encoded first, followed by the COMBINING GEMINATION MARK.

- In §9 Unicode Properties, The canonical combining class for VOWEL SIGN E and COMBINING GEMINATION MARK should be 230.
- In §4 Digits (and §9 Unicode Properties), explain why the Bidi_Class is AN (Arabic_Number), instead of EN (European_Number) or L (Left-To_Right).
- Discuss the use of the macron in the following (figure 24)



- Document the various uses of the squiggle in Garay in the proposal (i.e., based on comments from Jack Merrill, in vowel-initial words before the vowel mark; to host the vowel /e/ (or /é/) after voiced stops; to mark the +ATR vowels /ë/, /ó/, /é/ [likely inspired by the use of diacritics in the official orthography]; and used after the sign for /i/, when intended to mark the sound [ü]).

Recommendations: We recommend the UTC review this proposal and forward any comments, with those above, to the proposal author.

Note: This feedback has already been forwarded to the proposal author.

36b. Report

Document: [L2/17-322](#) Report on the Garay script 2017 (WG2 N4875) - Charles Riley

Comments: We reviewed this report from Charles Riley on his trip to West Africa. Having access to the pages of the Koran in the Garay script would be helpful. (No examples were contained in the latest version of the proposal, L2/16-069.)

Recommendations: We recommend the UTC review this FYI document at their leisure.

Note: This feedback has already been forwarded to the proposal author.



37. Bété

Document: [L2/17-323](#) Report on the Bété script 2017 (WG2 N4876) – Charles Riley

Comments: We reviewed this report on the Bété script, which contains a preliminary code chart.

The following comments were made:

- A clear case needs to be made for encoding this script. Are documents available?
- It appears that the same signs are used to indicate something else by doubling:

1E600		ba	ba	BETE SYLLABLE	BA
1E603		beu	bə	BETE SYLLABLE	BEU

- The proposal should distinguish the graphical elements of a syllabary, as opposed to a sequence that is mapped onto syllables.

Recommendations: We recommend the UTC review this FYI document at their leisure.

Note: This feedback has already been forwarded to the proposal author.

EUROPE

38. Latin

Document: [L2/17-437](#) Feedback on proposed Unifon letters – Marin Silva

Comments: We reviewed this document, which provided comments on a proposal from 2014, [L2/14-070](#), which was a revised proposal for Unifon letter additions by Michael Everson.

Two items from the original proposal ([L2/14-070](#)) were deemed “more urgent” in this document:

LATIN CAPITAL LETTER SMALL CAPITAL I WITH STROKE
LATIN SMALL LETTER CLOSED U.

In our opinion, the first, adding an uppercase version to U+1D7B LATIN SMALL CAPITAL LETTER I WITH STROKE, is not urgent. Are there any users actively asking for this character now? For the second character, LATIN SMALL LETTER CLOSED U, which is said to be used in Swedish dialectology, we recommend the author demonstrate its use in a full proposal, with examples.

Note also that encoding the second member of a case pair – just because one of the pair appears in Unifon – is not a strong argument, in our opinion.

This document also has comments on the name and glyph of LATIN LETTER TURNED-E R, suggestions on names for some letters, and an opinion on the unification of certain letters. However, unless a proposal is made for Unifon letters, preferably with support from users, no action needs to be taken.

Recommendations: We recommend UTC members review this document at their leisure.

SYMBOLS

39. Ancient Chinese Math Symbols

Document: [L2/17-219](#) Proposal to encode Ancient Chinese Mathematical Symbols (revised) - Kushim Jiang

Comments: We reviewed this revised proposal, and had the following comments:

- Provide the figure numbers beside the proposed characters.

- For COMBINING CHINESE PRIME and COMBINING CHINESE DOUBLE PRIME, use of U+0301 COMBINING ACUTE ACCENT and U+030B COMBINING DOUBLE ACUTE.
- For COMBINING CHINESE TRIPLE PRIME, a new character at U+20F1 in the Combining Diacritical Marks for Symbols block would be appropriate.
- Remove COMBINING ENCLOSING CUT-CORNER RECTANGLE. This mark is an editorial convention appearing in a math book. Such a use may be better handled via other means of representation, such as markup. In addition, getting such a mark to work could be difficult for implementers. Is this mark used outside math books? If two neighboring characters were mistakes, would the box enclose both characters? More information is needed.



- The CHINESE PLUS SIGN is well-justified by numerous examples. The representative glyph should have the horizontal shortened:

Current glyph: \perp

Example from figure 29:

Like the other operators (MINUS SIGN, DIFFERENTIAL SIGN, and INTEGRAL SIGN), it is larger than surrounding characters (see figure 29, above). As with the other large math operators, the Unicode Standard does not specify a particularly layout, but implementations are expected to follow accepted typographical conventions for layout (see 3.2.3. Large Operators in [UTR 25](#)).

- CHINESE MINUS SIGN similarly appears generally larger than surrounding characters, but in figure 26 it is the same size as POSITIVE DIFFERENCE SIGN, cf.

figure 26: vs. figure 15

Is this a pair of characters, or just one? Or is the sizing due to style of the headings in this particular work?

As with PLUS SIGN, the horizontal stroke in the representative glyph should be made shorter, so it matches the typical shape.

- The CHINESE PLUS-OR-MINUS SIGN, another large operator, is justified, based on figure 24:




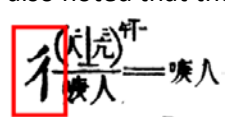

- CHINESE VARIANT PLUS SIGN and CHINESE VARIANT MINUS SIGN require more evidence. The evidence provided is in figure 25 (below). However, the book providing the justification seems to contain a different typographical style than the rest of the text. Provide more explanation on its

use. Unless a systematic distinction can be made between the variants and the MINUS SIGN and PLUS SIGN, we recommend they be removed.



- The CHINESE POSITIVE DIFFERENCE SIGN is shown on a heading in figure 27. The typed example on page 3 (below) suggests it is not a large operator – is that true?

$$5 \text{ 正 } 3 = 3 \text{ 正 } 5 = 2.$$

- For CHINESE FACTORIAL SIGN  we recommend a new character at U+20F2 with the proposed name COMBINING BOTTOM RIGHT CORNER (cf. 23CD/E/F), and an annotation, “used in Chinese.” The mirror image has been used in Western European math literature also for factorial (i.e., $\lfloor n$, see <http://mathworld.wolfram.com/Factorial.html>). (Note: The webpage <https://groups.google.com/forum/#!topic/sci.math/P20zFu2eebs> mentions that both were in vogue in England). (This alternate sign has not been proposed.)
- CHINESE LEFT CURLY BRACKET and CHINESE RIGHT CURLY BRACKET both require more examples showing their usage, beyond what is shown in figure 27. Until more evidence is provided, we recommend these be removed.
- CHINESE ELLIPSIS is shown in figure 27, which is a dictionary entry, but an example showing it in actual usage is required. Remove this character until more examples are provided.
- The CHINESE DIFFERENTIAL SIGN and CHINESE INTEGRAL SIGN appear to be acceptable. It was also noted that these two operators are larger than the neighboring characters. Cf. figure 29:
 
- This revision of the proposal includes the atomic (non-decomposed) circled symbols for inclusion in the Enclosed Ideographic Supplement block., which is the correct approach in our view.
- We recommend the whole set of Celestial Stems and 12 branches be proposed. Eight of the ten Celestial Stem characters are proposed here (U+1F252..U+1F259) but the listing should show U+1F25A and U+1F25B as “reserved” for the remaining two Celestial Stem characters (for “9” and “10”), when evidence is provided. The 12 branches are represented by U+1F270..1F27B, but a spot at U+1F27A should be reserved, when an example of the circled ideograph  is found.
- The other enclosed operators will require further investigation.

In sum, the following characters were deemed acceptable candidates:

U+20F1 COMBINING CHINESE TRIPLE PRIME
U+20F2 COMBINING BOTTOM RIGHT CORNER
CHINESE PLUS SIGN (with modification of glyph)
CHINESE MINUS SIGN (with modification of glyph)
CHINESE PLUS-OR-MINUS SIGN
CHINESE FACTORIAL SIGN
CHINESE DIFFERENTIAL SIGN
CHINESE INTEGRAL SIGN

Recommendation: We recommend the UTC review this document, send feedback to the author, and recommend the proposal be revised to just include the eligible characters (with evidence). For the characters deemed eligible, the UTC should consider whether CHINESE is appropriate in the characters' names.

Note: This feedback has already been forwarded to the proposal author.

-----NOT DISCUSSED-----

AFRICA

Egyptian Hieroglyphs

Document: [L2/18-165](#) Revised draft for the encoding of an extended Egyptian Hieroglyphs repertoire

SOUTH AND CENTRAL ASIA

Takri

Document: [L2/18-084](#) Proposal to encode the TAKRI VOWEL SIGN VOCALIC R – Srinidhi and Sridatta

EAST ASIA

CJK

Document: [L2/18-097](#) Request to add nine U-source ideographs - Jaemin Chung

NOTATIONAL/NUMBER SYSTEMS

Tally Marks

Document: [L2/18-088](#) Proposal to encode the two remaining tally mark systems proposed in L2/15-328 – Eduardo Marín Silva

SYMBOLS

Tachograph symbols

Document: [L2/18-090](#) On the encoding of tachograph symbols - Marius Spix

Feedback:

[L2/18-095](#) Feedback on proposal to encode tachograph symbols (L2/18-090) - Eduardo Marín Silva

OTHER (Properties)

a. Document: [L2/18-082](#) Extend the EquivalentUnifiedIdeograph property to three Suzhou numerals, etc – Eduardo Marín Silva

b. Document: [L2/18-120](#) New character property: NumberOfPeople – Eduardo Marín Silva