

TO: UTC
FROM: Deborah Anderson, Ken Whistler, Roozbeh Pournader, Andrew Glass, Laurentiu Iancu, Lisa Moore, Liang Hai, and Alolita Sharma
SUBJECT: Recommendations to UTC #151 May 2017 on Script Proposals
DATE: 5 May 2017

The Script ad hoc group met on 10 March and 28 April 2017 in order to review proposals. The following represents feedback on proposals that were discussed when the group met.

EUROPE

1. Palaeohispanic

Document: [L2/17-129](#) Proposal to encode the Palaeohispanic script – Joan Ferrer et al.

Comments: We discussed this proposal, which unifies “Southern Palaeohispanic” (proposed in [L2/15-119](#)) and “North-eastern Iberian” (proposed [L2/15-120](#)) into one script. (The 2015 proposals each provide extensive documentation and references.)

This revised proposal includes a chart with the representative glyphs (pp. 22-24 of the PDF). In our opinion, it would be easier to base the encoding on the graphemes, rather than on the phonetic values. (Note that Southern Iberian typically has a different sound value from Northern Iberian.)

For example, for TU and DU/TU, we recommend two characters be encoded,  and , and possibly a third,  if the horizontal line at the bottom is deemed to be significant. (If not, it could be considered a variant of .

Current chart:

	10221	tu												
	10222	du/ tu												

Suggested revised chart:

	10221	tu												
	10222	du/ tu												

A number of other, similar changes to the chart will need to be made, affecting: 10200/1, 10208/9, 10206/7, 10218, 1021D/E, 1021F/10220, 1022B, 1022C/D, 1022E/F. Some of the letter names will need to be changed as a result.

In addition, there was discussion on whether the the spelling “palaeo” should be retained, or “paleo” would be better. The first headword in the OED is palaeo-, so “Palaeohispanic” seems appropriate.

Recommendation: We recommend the UTC members review this document and forward comments to the proposal authors.

AFRICA

2. *Egyptian hieroglyphs – Quadrats*

Documents:

[L2/17-112](#) A method for encoding Egyptian quadrats in Unicode – Glass et al.

[L2/16-210](#) A system of control characters for Ancient Egyptian hieroglyphic text (revised) - Nederhof

Comments: We reviewed these documents, focusing primarily on Andrew Glasses’ contribution, (L2/17-112), though noting the revised version of L2/16-210 by Nederhof.

Glass’ “Method of encoding Egyptian quadrats” document reflects the result of ongoing discussions between Andrew Glass, Mark-Jan Nederhof, Bob Richmond, and other experts (named on page 1) which have taken place since a meeting of Egyptologists in Cambridge in July 2016 and the August 2016 UTC meeting.

The document L2/17-112 is exemplary as a complete, well-formed proposal: it clearly lays out the principles underlying the encoding model, discusses the various options with their pluses and minuses (pp. 4-5), provides examples that show the different control options (pp. 5-6), gives detailed examples of each control character and its relevant property information (pp. 7-12), and examines other relevant aspects, such as vertical layout (pp. 14-20) and OpenType strategies (pp. 21-25).

The document proposes 9 control characters (of which 2 have already been approved by the UTC).

The following comments arose during discussion:

- At the August 2016 UTC, there was a decision to investigate encoding options, including Polish notation. The encoding model adopted here uses binary controls with parentheses, which follows MdC syntax and avoids complex precedence rules.
- The OpenType model advocated is a hybrid solution (p. 23): that is, commonly occurring quadrats would be handled as precomposed glyphs (using ligatures), but a dynamic OpenType approach would be used for quadrats that are less common (p. 21)
- If the UTC accepts the proposed characters, the order of the two control characters as documented on the Pipeline will need to be changed (U+13430 would be the VERTICAL JOINER, and U+13431 the HORIZONTAL JOINER). The suggested order is based on operator precedence.
- Experts and UTC members should discuss when the STACK MIDDLE control character should be employed vs. atomic encoding, as this topic directly affects work on the extended Egyptian repertoire by Michel Suignard (see [L2/17-073](#))
- The proposal covers the majority of possible combinations of Egyptian hieroglyphs. Some cases not covered are mentioned in the section “Other insertions” on p. 10.
- The goal was to arrive at a solution that will be implementable, and can be extended to other scripts, such as Mayan hieroglyphs. Glass has offered to publish the OpenType details, which will make it easier for implementers.

Recommendation: We recommend the UTC discuss this proposal, which, in our view, represents a complete model that discusses all the issues.

3. *Loma*

Document: [L2/17-059](#) Towards an encoding of the Loma script - Everson

This is an FYI document; a linguist is providing additional information.

SOUTH ASIA

4. Bengali and Newa

Document: [L2/17-130](#) Comments on L2/16-322 and L2/16-383, Sandhi marks for Bengali and Newa - Anderson

Background documents:

[L2/16-322](#) Proposal to encode the SANDHI MARK for Bengali – Srinidhi and Sridatta

[L2/16-383](#) Proposal to encode the SANDHI MARK for Newa – Srinidhi and Sridatta

Comments: We reviewed the documents, which include the original proposals for *sandhi* marks for Newa and Bengali, and feedback from the proposal authors.

The authors disagreed with the decision at the January 2017 UTC to encode the two marks as combining *avagraha* signs (with gc=Mn, not gc=Po), and not as *sandhi* marks. The proposal authors report that the *sandhi* and *avagraha* signs are not the same, and need to be distinguished since they may occur in the same text.

The following points were made during the discussion:

- Because the examples are all from manuscripts and not printed texts, it is difficult to confirm that the placement of the *sandhi* marks is between syllables and not above the syllable; in some cases it is hard to tell, and placement can depend upon the visual elements.

Compare the following examples from Bengali *sandhi* proposal L2/16-322



vs.



and from Newa *sandhi* proposal L2/16-383



vs.



- Based on the decision to encode a Sharada *sandhi* with gc=Po ([L2/12-322](#)), it seems reasonable to expect that Newa and Bengali would be the same. In Sharada, the *sandhi* character is written after the syllable where *sandhi* occurs, and appears in a below-base position. It is distinct from the *avagraha*, which is a spacing mark that appears on the baseline.
- If there is evidence of a Bengali and Newa *avagraha* that is distinguished from the *sandhi* mark, a proposal can be prepared with the evidence. Such a proposal should take into account [L2/15-162](#) (“Script property of A8F1 COMBINING DEVANAGARI SIGN AVAGRAHA”), which shows an *avagraha*-like character in Bengali (for which it is suggested U+A8F1 might be used).

- Instead of focusing on the intended function of the sign, it might be more prudent to make a decision based on how the mark will best be handled for rendering engines. From this viewpoint, a combining character would be preferable. As a result, we recommend the proposal authors modify the general category properties to Mn;230;NSM;N;N. In order to be consistent, a change in the properties for U+111C9 SHARADA SANDHI MARK should also be made.

Recommendation: We recommend the UTC discuss this topic, and change the names of the following two characters (which were approved at the January 2017 UTC meeting):

	Current name	New name
U+09FE	COMBINING BENGALI SIGN AVAGRAHA	BENGALI SANDHI MARK
U+1145E	COMBINING NEWA SIGN AVAGRAHA	NEWA SANDHI MARK

We further recommend the UTC give an action to Ken Whistler and Laurentiu Iancu to create a proposal for changing the general category properties of U+111C9 SHARADA SANDHI MARK from Po to Mn (230;NSM; etc.) for Unicode 11.0.

5. Kannada

Documents:

[L2/17-041](#) Request to change glyphs of Kannada letters Vocalic L and Vocalic LL and their vowel signs – Srinidhi and Sridatta

[L2/17-127](#) Feedback and Response on L2/17-041: Request to change glyphs of Kannada letters Vocalic L and Vocalic LL and their vowel signs - Anderson

Comments: We reviewed these two documents: a proposal to change the glyphs for vocalic *l* and *ll* in Kannada and feedback to the proposal, with a response from the authors. The proposal feedback recommended the authors try to provide additional manuscript evidence for the glyph changes. The authors responded that it is difficult to access manuscripts, but they relied on secondary sources which drew from manuscripts.

The following comments arose during discussion:

- We noted that the only clean example of the dependent form of *l* and *ll* in L2/17-041 is in figure 6. Figure 8 has a dependent form of vocalic *l*, but it is unclear if it contains a vertical stroke.
- Based on the evidence provided, the proposed independent glyphs shapes for vocalic *l* and *ll* are strong. However, it is unclear whether the vertical line is a critical component of the dependent form. Without additional evidence, one could assume the dependent forms are subjoined versions of the independent forms.

Recommendation: We recommend the UTC approve the proposed glyph changes for the independent forms of vocalic *l* and *ll* (U+0C8C and U+0CE1) and discuss whether to accept the new glyphs for the dependent forms of vocalic *l* and *ll* (U+0CE2 and U+0CE3). If approved, the authors should provide a font whose glyphs match the ductus of the characters in the current Kannada chart.

6. Nandinagari

Documents:

[L2/17-119](#) Towards an encoding model for Nandinagari conjuncts - Pandey

[L2/16-310](#) Revised proposal to encode Nandinagari - Pandey

Comments: We reviewed the above documents. The following summarizes the discussion.

Encoding model

- The main focus of discussion was on the encoding model, specifically how to handle “touching forms”. “Touching forms” are a way to represent consonant clusters that touch one another, without modification, or whose headlines touch.

For example:

Regular ligature: Touching form:

dda ३ ३३

Based on the variety of forms shown on pp. 40-47 of the proposal (L2/16-310), such as for *ddha* below, the ad hoc concluded that touching forms appear to be just other conjuncts and don't apparently need to be distinguished in plain text.

<i>ddha</i>	३५,
	३५,
	३५,
	३५,

If, however, touching forms need to be distinguished in plain text, then a proposal should be provided and the topic can be discussed.

Ardhavisarga

- There was some discussion about the *ardhavisarga*, which is used to indicate both *jihvamuliya* and *upadhmaiya*. Can the *ardhavisarga* take marks or touch other letters? (Cf. Kannada *jihvamuliya* on p. 498 of *TUS*, which can take a dependent vowel sign.) Would use of the Vedic *ardhavisarga* U+1CF2 be acceptable in place of a script-specific Nandinagari *ardhavisarga*? (Note that VEDIC SIGN ARDHAVISARGA is Mc, but the Nandinagari character is Lo, as are Kannada *jihvamuliya* and *upadhmaiya*.)

Vedic

- Check with experts whether U+11BD8 NANDINAGARI SIGN VEDIC ANUSVARA and U+11BD9 NANDINAGARI SIGN DOUBLE VEDIC ANUSVARA are used in other scripts (and hence would also be eligible for inclusion into the Vedic Extension block, with an entry in ScriptExtensions.txt)

Other notes

- Mention that the styles of Nandinagari may vary, depending upon calligraphic style (cf. long swashes in figure 15). This may affect Indic properties, such as Indic Positional Category (cf. dependent vowel *i* in figure 15).

Recommendation: We recommend the UTC invite a revised proposal from the author, incorporating the comments above.

7. Tamil

Documents:

[L2/17-069](#) Tamil Fractions and special symbols in Unicode – Gov't of TN

[L2/17-114](#) Comment on L2/17-069: Tamil Fractions and Symbols – Ganesan

Comments: We reviewed this proposal for 51 Tamil historical fractions and symbols. The proposal was submitted by the Tamil Virtual Academy under the Government of Tamil Nadu, and represents expert feedback. The proposal has seen several rounds of review by the UTC and experts in Tamil Nadu. We laud those responsible for coming to consensus and submitting a final version. (We also note the approval by Ganesan in L2/17-114.)

The UTC had earlier approved Tamil historic signs in May 2013, but modified some names in February 2015. As mentioned in the proposal, the characters had been included in the CD2 of the fifth edition of ISO 10646 (see L2/15-339), but were later removed based on ballot comments from the Government of India (L2/16-091).

A few comments were made during discussion:

- The representative glyphs are based largely on historical shapes, though it was noted that later glyph shapes can be represented in fonts
- The recommendation to disunify some characters which were not disunified earlier is discussed in section 3 (p. 6 of PDF)
- The proposal lists the document trail, but it was noted there was also a script ad hoc report also on the topic ([L2/16-155](#) with an accompanying [spreadsheet](#)).

Recommendation: We recommend that the UTC resynchronize its approval of Tamil fraction and other symbols with the details for code points and character names for the 51 Tamil characters shown in L2/17-069, after discussion.

8. Vedic

Document: [L2/17-117](#) Proposal to encode a nasal character in Vedic Extensions – Pandey

Comments: We reviewed this proposal for one Vedic character.

To represent a Vedic *anusvara*, the Nandinagari script proposal had recommended using the Vedic character U+1CE9 VEDIC SIGN ANUSVARA ANTARGOMUKHA. However, the Nandinagari proposal had proposed a script-specific character for the *double anusvara antargomukha*, a sign that appears in Vedic texts and is made up of two *anusvara* stacked together. The script ad hoc group had recommended in [L2/17-037](#) that this character be moved from Nandinagari into the Vedic Extensions block.

This short proposal includes an example and properties. The next available code point in the Vedic Extensions block is U+1CFA. The proposal should specify that the script=Common and Script Extensions property should be Nandinagari. Also, the Indic Syllabic category should note that the character participates in syllables.

Recommendation: We recommend the UTC approve the character U+1CFA VEDIC SIGN DOUBLE ANUSVARA ANTARGOMUKHA, making the updates to the proposal as suggested above.

9. *Wancho*

Document: [L2/17-067](#) Proposal to encode the Wancho script (WG2 N4787) - Everson

Comments: We reviewed this proposal for Wancho, a simple script with no conjuncts. This proposal, which provides details and examples, builds off [L2/17-042](#), which was a very barebones preliminary proposal.

The following comments were raised during discussion:

- The script is still under development.
 - Tones in this proposal are indicated by doubling or tripling the letter. Ongoing discussion is taking place about marking tone with via a different mechanism (i.e., such as by a character).
 - An inherent vowel occurs when two consonants come together in certain contexts. There is discussion about creating a new character to represent this vowel.
 - In our view, it is still possible to propose the script for encoding, and add characters later (which would require proof of use in any event). It appears that the current repertoire is being taught in the schools.
- In the proposal, provide information on the range of dialects (/languages) the script could cover (i.e., Upper, Middle, and Lower Wancho).

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

SOUTHEAST ASIA

10. *Nyiakeng Puachue Hmong*

Document: [L2/17-002r3](#) Proposal to encode the Nyiakeng Puachue Hmong script – Everson

Comments: We reviewed this revised proposal, which has addressed comments from the previous ad hoc report ([L2/17-037](#)).

The proposal now includes an example of the logogram NYAJ (in figure 12), indicating money. Understandable names of the determinatives and tone marks have been provided, and the XW XW character is cross-referenced to similar character in Lao and Thai. The new name of the script is a slight improvement from the earlier names (Nyiakengpuachue, and then Hniakeng Puachue); the proposal author was not able to accommodate a more simplified name. The sole example of U+1E14F CIRCLED CA is in figure 13. A better example is still needed for this mark of ownership.

Recommendations: We recommend the UTC review this proposal, and approve the script for encoding, after discussion about U+1E14F CIRCLED CA.

MIDDLE EAST

11. *Arabic*

Al-Dani Quranic Marks

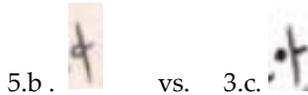
Document: [L2/17-068](#) Suggestions on Hamza Al-Dani Quranic Marks proposition - Lazrek

Comments: We reviewed this document, which provides additional new information, but also reflects some confusion about the Arabic encoding model in Unicode.

This proposal does not advocate atomic encoding of characters, such as those recommended in the [January script ad hoc report](#), but instead prefers encoding base characters with combining characters. Lazrek's approach to the encoding model will be further described in a forthcoming proposal for *hamzas*. It was noted during the ad hoc discussion that generally Arabic marks that attach to a base become part of the character (rather than being understood as a separate combining character).

The following comments were noted:

- In 5a, b, c, the circle is clearly connected to the alef, but in 3c it is disconnected (Also shown on page 3, in examples, 3, 4 and 5). Are these the same? Could the hollow dot and the filled dot¹ be considered the same?



It was noted that for the Koranic marks, the tendency in Unicode has been towards a visual encoding. If the circle and filled dot both occur and are deemed different enough, then both a hollow and filled dot may need to be encoded.

- Provide examples of (a) top two characters on page 9, and (b) the third and fourth characters from the bottom on p. 10 (considered variants by Lazrek) (see below).

Page 9:



Page 10:



- The author should check the links in his proposal.

Recommendation: We recommend the UTC review the documents and relay their comments, including the comments above.

EAST ASIA

12. Hangul

Documents:

[L2/17-125](#) Preliminary Proposal for Encoding New Hangul – S. and S. Oh

[L2/17-126](#) Comments on "Preliminary Proposal for Encoding (new hangul jamo for some English consonant sounds)" – Jaemin Chung

Comments: We reviewed this preliminary proposal, which proposes new Hangul Jamo for English consonants. In our opinion, the proposal needs to show that the proposed characters are in common use. The document by Jaemin Chung also lays out other specific comments, including the need for images of the proposed characters.

¹ In the recommendations from the January 2017 ad hoc, L2/17-037, the recommendation was for an ARABIC ROUNDED STOP WITH FILLED CENTRE, a non-combining punctuation mark that would appear to the left of the stem of the *alef*.

Recommendation: We recommend the UTC review the documents and relay their comments, including the feedback in L2/17-126 and the comments above.

CENTRAL ASIA

13. *Jurchen and Khitan Small Script*

Document: [L2/17-107](#) Towards an encoding of the Jurchen script and implications for Khitan Small Script – West, Everson, and Zaytsev

Comments: We reviewed this document, which proposes a new “Ideographic Radicals” block for Jurchen and Khitan Small Script radicals.

The following observations were made:

- Radicals are needed primarily for lexicographic purposes; in Khitan Small Script, the radicals are not intrinsic parts of characters (compared to Tangut), and they do not appear widely in modern texts.
- Khitan Small Script has 20 components used as radicals; these are currently incorporated as part of a single Khitan Small Script block repertoire in the ISO PDAM 1.2 ballot. Twelve radicals do not participate in clustering.
- Jurchen has a total of 50 radicals; the script is fairly well understood, and a new proposal is being prepared by Andrew West for the Sept. 2017 WG2 meeting
- Details regarding radicals for Khitan Small Script and Jurchen:
 - Khitan Small Script and Jurchen share 15 radicals between themselves
 - The total number of Khitan Small Script and Jurchen-only radicals is 55; of these, 29 are shared with CJK or Kangxi radicals
 - Of the 20 Khitan Small Script radicals, 9 are shared with CJK or Kangxi radicals
 - Amongst the 35 radicals only occurring in Jurchen, 20 are shared with CJK or Kangxi.
 - Jurchen shares 1 radical with Khitan Large Script (which also has no full proposal yet)

Discussion touched on the following points:

- The document should address why the Khitan Small Script radicals should be disunified from the Khitan Small Script block of ideographs. What is the implementation argument why this is needed?
- The Jurchen proposal should discuss the decision to encode it as a separate script, and to not unify it with Han (though there is significant overlap).
- A separate step would be to provide a clear argument to disunify radicals from the main set of Jurchen ideographs. Is there an implementation requirement for maintaining a distinction of radicals?

Recommendation: We recommend the UTC discuss this document. The proposed change would affect Khitan Small Script characters approved by the UTC and currently on the Pipeline.

14. *Mongolian*

Document: [L2/17-128](#) Modify Standardizedvariants.txt file in the Mongolian/ToDo Block U+1800 - U+185C Phase II -- Munkhdelger Dorji, et al

Comments: We reviewed this document, which proposes changes to the Standardizedvariants.txt file for Mongolian and ToDo, but not Sibe or Manchu. The proposed changes have been generally accepted by the Mongolian community. Some examples are provided.

There was a production error on 1820 180B which will be fixed in a revised version of this document.

Recommendation: We recommend the UTC discuss this document, but hold off making any decisions until the encoding model for Mongolian is worked out in more detail.

AMERICAS

15. *Mayan Numerals*

a. **Mayan numerals script proposal**

Document: [L2/16-264](#) Mayan Numerals - Quinn

Comments: We briefly reviewed this document, which incorporated changes recommended at the November 2016 UTC (#149), as had been captured in the script recommendations ([L2/16-342](#)).

Recommendations: We recommend the UTC note this document, but no action is required.

b. **Mayan Numerals layout**

Document: [L2/17-110](#) Layout and reading order in Mayan historical texts – Pallan and Anderson

Comments: We reviewed this document, which provides details on how Mayan numbers work in the codices and inscriptions. As was discussed in the script ad hoc recommendations [L2/16-342](#), the vertical orientation should be U (“upright”). It is advisable to treat the individual numerals as symbols by themselves. Beyond that, the actual orientation will depend upon the application for layout.

There was a question posed about one glyph: Does the image on the left indicate ‘11’? (If so, it varies from the more usual layout, as on the right)



(According to C. Pallan, the two vertical lines with a dot on top are not ‘11’, but instead are interpreted as a death sign, specifically a disembodied eye.)

Recommendations: We recommend the UTC note this document.

SYMBOLS

16. ISO/IEC keyboard symbols

a. Document: L2/17-047 Proposal to encode ISO/IEC 9995 symbols used on German keyboards
– Pentzlin

Comments: We reviewed this proposal, which requests for encoding a small subset of the characters drawn from the larger proposal L2/17-072 (12 characters in L2/17-047 vs. 108 characters + 2 named sequences in L2/17-072). All the characters in the proposal derive from ISO/IEC 9995-7:2009 and its Amendment 1 (2012) and from ISO/IEC 9995-10:2013.

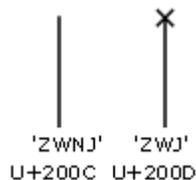
The following were noted:

- 10 of the 12 characters were proposed in 2010 (L2/10-351), 2011 (L2/11-074 and L2/11-133) and 2012 (L2/12-302). Two new characters appear in the 2017 version:
WHITE HORIZONTAL NARROW RECTANGLE (deadkey space mark)
WHITE VERTICAL NARROW RECTANGLE (cap height marker, identifying modifier letters on keycaps)
- As mentioned in the proposal, the characters were not approved earlier because “actual use” of such symbols was required. The current document includes an image showing a portion of a German keyboard and a snippet of text from the German standard DIN 2137-1:2012-06. The current proposal mentions that “some” of the symbols are in use. However, appearance of a graphic symbol in an ISO standard is not a guarantee that a character need be encoded to represent it. The USNB feedback sent in 2011 (L2/11-078) still rings true:

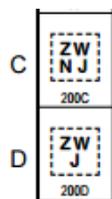
We would welcome attestation and analysis of each symbol in wide use as text characters, and a full proposal along these lines. It would also be useful if the proposal documented keyboard symbols that are in widespread use for functions that correspond to those covered by the repertoire of symbols in the proposal.

- Questions to answer:
 - Are the characters in widespread use? If so, provide evidence in text, such as hardware or software manuals.
 - Why are images not sufficient?
 - Show the variety of symbols being used today.
For example, Microsoft has different glyphs for ZWNJ and ZWJ (<https://www.microsoft.com/typography/otfntdev/glyphs.htm>). If the Microsoft glyphs are widely used, encoding rarely used glyphs could be counter-productive. How common are the dotted box glyphs for ZWNJ and ZWJ?

Microsoft:



Current glyphs:



Proposed:



Compare also the Android ZWNJ (from <http://images.farnet.ir/2014/06/Android-L-Persian-Keyboard.jpg>):



- The menu symbol  had been seen by a member of the ad hoc, and in a casual search it was found on a keyboard (<http://madoxweb.com/how-to-control-windows-with-only-a-keyboard/>), so if documented widespread use is provided, it may be eligible for encoding.

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

b. Document: [L2/17-072](#) Proposal to incorporate the symbols of ISO/IEC 9995-7:2009 and its Amendment 1 and of ISO/IEC 9995-10:2013 into the UCS – SC35

Comments: We briefly reviewed this large proposal. See the comments above, for L2/17-047.

Recommendations: We recommend the UTC review this proposal and respond to the submitters, requesting additional plain text evidence, and address the questions listed above.

17. MR Sign

Document: [L2/17-066](#) MR Sign Proposal to encode the Marca Registrada sign – Eduardo Marín Silva

Comments: We reviewed this proposal, which provided several examples. At this point, we do not see a need to distinguish a circled MR, but it could be proposed if a need is shown.

Recommendations: We recommend the UTC accept the character U+1F16C RAISED MR SIGN.

Currency symbols

18. TOMAN sign

Document: [L2/17-060](#) Proposal to encode Iranian Currency Sign TOMAN to the UCS – Toman O Rial

Comments: We reviewed this proposal for a monetary unit approved by the Iranian government in 2016.² Although the word “*toman*” is used in speech and writing, it does not necessarily mean that a currency sign exists. The sole evidence in the proposal is an old bank note which stylistically spells out the word, but does not represent a symbol. In comparison, the RIAL symbol, though rare, was found on typewriters and in the national standard (cf. [L2/01-354](#)).

Recommendations: We recommend the UTC not accept this character. If evidence is provided of this as a symbol, the submitter is invited to re-submit a proposal.

19. Vexillology

² According to Wikipedia, it is still awaiting approval from the Iranian parliament.

Document: [L2/17-089](#) Preliminary proposal to encode Vexillology Symbols in Unicode -- Anshuman Pandey

Comments: We briefly reviewed this preliminary proposal, which proposes a set of symbols that relay information on the use and features of national flags.

The first four columns of characters have a six-part grid pattern (++) . Dots (or x's) in the grid indicate use of the flag for land (top row) and sea (bottom row), and private, public, and military (in the three columns). The first four columns of symbols have been adopted by Fédération internationale des associations vexillologiques (FIAV). The last two columns contain 19 flag typology symbols (U+1FA40..U+1FA52), which have never officially adopted by FIAV, according to the nameslist.

While the iconography to represent flags is important, the author should make a case why it requires character encoding. Examples of plain text usage should also be provided. If evidence is strong, the block should be added to the Roadmap.

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

Numbers

20. *Roman numerals*

Document: [L2/17-092](#) Proposal to add Roman Numerals Alternate Number Fours -Eduardo Marín Silva

Comments: We reviewed this proposal, which requested two characters, an upper- and lowercase Roman Numeral Alternate Four (IIII and iiii). The rationale for encoding these forms is based in part on the CJK compatibility characters for Roman numerals (U+2160..U+217F), but this is not, in our opinion, strong evidence.

As noted in the last example on the last page, CCCC is handled as four C characters. We think a similar approach would apply here as well: users should just use four I or i characters. (Cf. Chapter 22 of *TUS*, "For most purposes, it is preferable to compose the Roman numerals from sequences of the appropriate Latin letters.")

Recommendations: We recommend the UTC briefly review this proposal and forward feedback (including that above) to the proposal author.

21. *Counting rods*

Document: [L2/17-085](#) Proposal to add 6 standardized variation sequences for counting rods -- Eduardo Marín Silva

Comments: We reviewed this proposal, which proposes six standardized variation sequences for the Southern Song form of the counting rods. In our opinion, use of Variation Sequences is not an appropriate mechanism. Because the Southern Song symbols are very different from the Counting Rod Numerals, different characters should be proposed. The author should also provide more examples, similar to the proposal for tally marks by Ken Lunde ([L2/16-046](#) and its earlier version, [L2/15-328](#))

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

22. Chess symbols

a. Heterodox chess symbols (Fairy chess)

Document: [L2/17-034](#) Revised proposal to encode heterodox chess symbols in the UCS – Wallace and Everson

Comments: We reviewed this revised proposal for 84 characters. An earlier version of this document was discussed at the January UTC, but it has been updated with examples of all the glyphs. The proposal builds off L2/16-293 by Garth Wallace.

The following comments were made:

- In general, it is advisable not to rely only on online evidence. The earlier ad hoc report (L2/16-342) had recommended: “Provide more than one example of each character (hence, include another example besides the Wikipedia article), drawing on print material, or highly formatted text on the Web.” Relying solely on a glyph in a font is not considered strong evidence.
- Provide additional examples of U+1FA4D NEUTRAL CHESS EQUIHOPPER ROTATED NINETY DEGREES. The current evidence is based on figure 43, which is from a LaTeX Diagram manual. Additional examples from sources would ensure the characters are not ghost characters.
- The glyph shapes for the hybrid chess symbols varies (see Tables 1, 2, and 3, pp. 29-30). Provide evidence of usage for the glyph shapes for the hybrid chess symbols (outside of Wikipedia):
 - WHITE CHESS KNIGHT-QUEEN
 - WHITE CHESS KNIGHT-ROOK
 - WHITE CHESS KNIGHT-BISHOP
 - BLACK CHESS KNIGHT-QUEEN
 - BLACK CHESS KNIGHT-ROOK
 - BLACK CHESS KNIGHT-BISHOP

Recommendations: We recommend the UTC discuss this proposal.

b. Chess notation

Document: [L2/17-077](#) Proposal to add standardized variation sequences for chess notation (WG2 N4793)

Comments: We reviewed this proposal, which proposes a way to standardize how chess fonts are encoded, via 34 variation sequences. Sequences with one variation selector, VS1, would indicate pieces on a white square, whereas those with VS2 could be used to indicate pieces on a black square. The proposed sequences include two sequences to indicate white and black squares. Sequences for draughts are also included.

According to the proposal author, chess fonts have not used Unicode chess characters because no standardized mechanism is available to create chess diagrams (so chess pieces can be displayed on black and white squares).

The following comments were noted:

- An alternative but preferable approach would be to recommend the chess community convert their fonts to Unicode, a change that would not require rendering engines be changed, and then come to the UTC with their outstanding issues.

- If the interchange problem is for boards, then the correct level of approach would be to define a Chess Markup Language, which would understand and convey the board structure (and can make use of the Unicode chess characters as the character elements). A Chess ML would, however, likely require a specialized renderer. (Note that the Chinese NB had earlier proposed a way to describe the Chinese chess board based on box-drawing characters in [L2/10-368](#).)

Recommendations: We recommend the UTC discuss this proposal, but not approve it.

The following proposals and documents were in the document registry when the script ad hoc met, but not discussed:

EUROPE

Latin

[L2/17-076](#) Proposal for the encoding of an Egyptological YOD - Suignard

AFRICA

Egyptian hieroglyphs

[L2/17-073](#) New draft for the encoding of an extended Egyptian Hieroglyphs repertoire - Suignard

SOUTH ASIA

Indic (general)

[L2/17-098](#) Request for editorial updates to Indic scripts –Srinidhi A, Sridatta A

Devanagari Numbers

[L2/17-116](#) Clarification on Devanagari letter numerals – Srinidhi

Reference documents:

[L2/17-010](#) Preliminary proposal to encode Devanagari letter numerals

[L2/17-040](#) Feedback on preliminary proposal for Devanagari letter numerals – Sharma

SOUTHEAST ASIA

[L2/17-106](#) Proposal to Encode Lao Characters for Pali – Vinodh Rajan et al.

CENTRAL ASIA

Khwarezmian

[L2/17-054](#) Preliminary proposal to encode the Khwarezmian script - Pandey

Mongolian Suffix Connector

[L2/17-052](#) Comments on L2/17-036 Mongolian Suffix Connector – Liang Hai

[L2/17-123](#) Comments on L2/17-036 (MONGOLIAN SUFFIX CONNECTOR) – Greg Eck

Background doc:

[L2/17-036](#) Encode Mongolian Suffix Connector (U+180F) To Replace Narrow Non-Breaking Space (U+202F) – Eck et al.

Mongolian

[L2/17-063](#) Proposal to Encode One Sibe Letter and One Mongolian Format Control, and Add Ten Variants to Three Mongolian Letters – Xudong Ma, et al

L2/17-064 Proposal to Change Glyphs of Mongolian Letter Manchu Ka – Xudong Ma, et al
L2/17-065 Proposal to Correct the Descriptions of Mongolian Letter Manchu Ka and Manchu Vowel Harmony – Xudong Ma, et al
L2/17-115 Glyph corrections for 11 Ali Gali final forms – Weizhe
L2/17-124 DS01 MONGOLIAN BASE FORMS, POSITIONAL FORMS, & VARIANT FORMS – Eck et al.

MIDDLE EAST

Elymaic

L2/17-055 Preliminary proposal to encode the Elymaic script - Pandey

EAST ASIAN

Hangul

L2/17-080 Informative document about three pre-Unicode-2.0 modern hangul syllables - Jaemin Chung

L2/17-081 Proposal to add an informative note to U+3164 HANGUL FILLER - Jaemin Chung

L2/17-090 Proposal to add informative notes and cross-reference to U+F92C and U+F9B8 - Jaemin Chung