TO: UTC
FROM: Deborah Anderson, Ken Whistler, Rick McGowan, Roozbeh Pournader, Anshuman Pandey, and Andrew Glass
SUBJECT: Recommendations to UTC #143 May 2015 on Script Proposals
DATE: 3 May 2015

The recommendations below are based on documents available to the members of this group at the time they met, and do not include documents submitted later to the document registry.

SOUTH ASIA
Indic
1. Bengali
Documents:
L2/15-091 Bengali letter UA used in Kokborok - Sanghmitra Sahu

Discussion: We reviewed this document, which provides documentation on the Bengali letter UA, which is used in the Kokborok language of Tripura, India, and surrounding areas.

Although the document included a proposal summary form requesting one character, the letter can already be written as a sequence U+0989 BENGALI LETTER U followed by U+09BE BENGALI VOWEL SIGN AA, and is supported on recent OSes: উা

Recommendations: We recommend the UTC respond to the author, thanking her for the documentation, and relay that BENGALI LETTER UA can already be handled as a sequence of <U+0989, U+09BE> on recent OSes.

Like the earlier proposal by the same author (L2/14-304), we recommend the Editorial Committee add text to the Bengali section of 8.0 alongside that for the letter AW, i.e., mention that the UA form exists, it can be represented by the sequence <U+0989, U+09BE>, and is an exception to the rule that independent vowels are separately encoded.

2. Gondi
a) Masaram Gondi
Document: L2/15-090 Proposal to Encode the Masaram Gondi Script – Pandey
Discussion document: L2/15-112 GONDI and GUNJALA GONDI CHARACTER NAMES – Vowels EE and OO – Ganesan

Discussion: We reviewed the proposal, which is a revision of L2/15-005. The new version incorporates changes based on comments from UTC members, and has been reviewed by two experts.

The key changes to this proposal (compared to the earlier version seen at the February 2015 UTC) include:
- adding "Masaram" to the name (hence making it distinct from another script for the same language, see Gunjala Gondi, below)
- separately encoding REPHA and RA-KARA (option 4 on p. 11), instead of using sequences with a virama

- adopting a model that contains a pure killer (HALANTA), an invisible stacker (VIRAMA), and a cluster final RA (RA-KARA)
- encoding as atomic characters the clusters KSSA, JNYA, and TRA, since conjuncts in Masaram Gondi are written with linear sequences of half-forms, not as ligatures.

The following points were raised during the discussion:
- the proposal should explicitly state on the top of page 9 that vowels come after the RA-KARA (i.e., the order is consonant - RA-KARA – vowel sign), and that the virama is not to be used after RA-KARA
- add script extensions information for danda
- change the CCC for HALANTA from 0 to 9 (the class for all viramas)
- request the Roadmap Committee to change the name for this script from "Gondi" to "Masaram Gondi"
- make adjustments as needed to 5.3 Syllabic Categories, pending discussion on where to put RA-KARA and REPHA.

The discussion document L2/15-112 by Ganesan requests the currently proposed Masaram Gondi vowel letters E and O be moved down one code point and renamed EE and OO, based on correspondences with other South Asian scripts and the phonology of other Dravidian languages. Based on Pandey's comment (page 3 of L2/15-090) that the script does not have letters or dependent signs for the Dravidian long vowels /e:/ and /o:/, we feel the current encoding should stand as proposed. If the script later adds characters for these letters, they can be encoded and given an appropriate name.

We feel this proposal is mature and the script is ready for encoding.

Recommendations: We recommend the UTC review the proposal, noting comments into account, and approve Masaram Gondi for encoding (with revisions that arise during discussion).

b) Gunjala Gondi
Document: L2/15-086 Preliminary Proposal to Encode the Gunjala Gondi Script - Pandey
Discussion document: L2/15-112 GONDI and GUNJALA GONDI CHARACTER NAMES – Vowels EE and OO – Ganesan

Discussion: We reviewed this preliminary proposal, which is for a re-discovered script for the Gondi language. The script was inspired by Modi and has received media attention. The proposal author expects to get feedback from the user community in the near future.

The script ad hoc had the following comments:
- Add script extensions information on the dandas
- In figures 9 and 10, mention Telugu
- Add the ISO 639-3 language code for Gondi
- Provide a comparison chart between Gunjala Gondi and Modi
- Identify the areas where Gunjala Gondi is used (versus where Masaram Gondi is used)
- Add text that the script does not currently distinguish Dravidian /e:/ and /o:/

Ganesan's document L2/15-112 requests two code points be reserved for E (at U+11DB6) and for O (at U+11DB9), and that the currently proposed vowel letters E and O be renamed EE and OO and moved (to

U+11DB7 for EE and U+11DBA for OO). As for Masaram Gondi, the rationale for this request is based on correspondences with other South Asian scripts (such as Tamil, Telugu, Kannada, and Grantha) and the phonology of other Dravidian languages. Although Pandey's Gunjala proposal does not mention that the script has no distinction for the Dravidian long vowels /e:/ and /o:/, and we feel the proposed encoding is acceptable. If the script later adds characters for these letters, they can be encoded and given an appropriate name.

Recommendations: We recommend the UTC review this preliminary proposal, and send feedback to the author.

3. Grantha
Document:
L2/15-085 Representing the fractions 1/4 and 1/20 in Grantha text – Sharma

Discussion: We reviewed this document concerning how two fractions in Grantha should be represented, given that the Tamil and Grantha numbers are unified. The fraction for 1/4 in Tamil should be represented with the glyphically identical U+0BB5 வ TAMIL LETTER VA and 1/20 with U+0BAA ப TAMIL LETTER PA (see L2/13-047), but how should the same fractions in Grantha be handled?

Shriramana recommends Grantha fractions be represented with the Tamil characters, since the glyph for Grantha LETTER PA (U+1132A) is different. We find Shriramana's recommendation to be reasonable.

Recommendations: We recommend the UTC review the document and make a decision.
(Note: There is a typo in L2/15-085: In the middle of the page VA is cited in parentheses with the code point "0BAA" and PA with "0BB5". These should be reversed, since VA is U+0BB5 and PA is U+0BAA.)

4. Gujarati
Document: L2/15-103 Revised Proposal to Encode Gujarati Signs for the Transliteration of Arabic – Pandey

Discussion: We reviewed this document, which is a revision of L2/14-131. The proposed characters are used to transliterate the Arabic script into Gujarati, and are used by the Ismaili Khoja communities.

The ad hoc discussion focused primarily on the SUKUN and the nuktas, since these characters reflect the representation of Arabic sounds being applied to a non-Arabic script.

The author will conduct further research into other letters that might appear as superscripts (besides LETTER A and LETTER MA, discussed in section 6), and may propose superscript characters at a later point.

Recommendations: We recommend the UTC encode these marks, but discuss Indic Syllabic categories on page 6, and whether "7" is the appropriate CCC value for the GUJARATI SIGN THREE-DOT NUKTA ABOVE and GUJARATI SIGN CIRCLE NUKTA ABOVE.

5. Tamil
Document: L2/15-115 Tamil Symbols Names and Annotations – Comment on L2/15-078 & L2/15-079 - Ganesan

Discussion: We reviewed this document.

In Part I of Section 2.0, Names and Annotations, Ganesan advocates four Tamil names in ழ be spelled with ZH, and not L. (The USNB ballot comments on PDAM 2.3 had the reverse: we requested the names be spelled with L and not ZH.) Part II of the document concerns normative aliases for the Tamil symbols.

A meeting took place on 29 April with interested stakeholders to discuss the Tamil symbols, and another meeting on the topic is scheduled for later.

Recommendations: The UTC should consider whether to make comments on DAM2 on the spelling of four characters:
> 11FD0 TAMIL FRACTION DOWNSCALING FACTOR KIIL > TAMIL FRACTION DOWNSCALING FACTOR KIIZH
> 11FD3 TAMIL SIGN AALAAKKU > TAMIL SIGN AAZHAAKKU
> 11FD5 TAMIL SIGN MUUVULAKKU > TAMIL SIGN MUUVUZHAKKU
> 11FE0 TAMIL SIGN KULI > TAMIL SIGN KUZHI

Regarding Part II, we suggest the UTC remand the name aliases to the Editorial Committee.

6. Telugu
Document: L2/15-082 Need for representing often used Telugu symbols - Challa Anilkumar

Discussion: This document states that certain combinations of Telugu letters do not show up well or properly. The two letters identified as being problematic in combinations are the "historic phonetic variants" U+0C58 TELUGU LETTER TSA and U+0C59 TELUGU LETTER DZA. These two characters were added in Unicode 5.1 (2008).

Upon testing, members of the script ad hoc determined that the combinations are valid on Android, but some combinations are rendered incorrectly. The Windows shaping engine doesn't prohibit combinations with these letters, but again the font fails. Hence, in our view, the problem is not one of Unicode, but due to the font (and possibly using an outdated rendering engine).

Recommendations: We recommend the UTC respond to the author, recommending the author be sure the most recent rendering engine is used and to notify Telugu font providers that combinations with U+0C58 TELUGU LETTER TSA and U+0C59 TELUGU LETTER DZA are not producing the expected result.

7. Vedic
Document: L2/15-101 Encoding of Vedic characters used in non-Devanagari scripts - Srinidhi
Discussion document: L2/15-113 Comments on L2/15-101 on non-Devanagari Vedic characters - Sharma

Discussion: We reviewed the preliminary document by Srinidhi on Vedic characters in non-Devanagari scripts, and the feedback document by Sharma.

As presented, nothing contained in the document from Srinidhi (L2/15-101) is immediately actionable, though we note that a comprehensive review of Vedic characters across the Indic scripts is a desideratum.

Below are a few comments from the ad hoc committee:
- The middle dot character (page 2), if proposed later, would require evidence that it belongs to the syllabic structure. Is there another encoded character that could be used? What is its function? Is this just a one-off orthography or is it a very common character?
- The Samavedic symbols in Bengali (pp. 14-15) are similar to the svara marks encoded in the Devanagari Extended block and seem reasonable to encode, if full attestation is provided.
- An overall recommendation would be to encourage Srinidhi (and other specialists) to identify Vedic symbols that occur in various scripts, and to submit feedback if they are not included in ScriptExtensions.txt (http://www.unicode.org/Public/UCD/latest/ucd/ScriptExtensions.txt).

Recommendations: We recommend the UTC members solicit feedback from interested parties, who can submit proposals for those characters deemed eligible for encoding.

CENTRAL ASIA

8. Old Sogdian
Documents: L2/15-089 Preliminary Proposal to Encode the Old Sogdian Script - Pandey

Discussion: We reviewed this preliminary proposal, which incorporates feedback from the expert Nicholas Sims-Williams.

Recommendations: We recommend the UTC review the proposal and send the author feedback.

9. Soyombo
Documents: L2/15-094 Response to Mongolian and Japanese comments on Soyombo and Zanabazar Square – Pandey
Related document: L2/15-009 Comments on Proposals of Zanabazar Square and Soyombo Script from Mongolian Experts (WG2 N4653) – Toshiya Suzuki, et al

Discussion: We reviewed this document, which was a response to the comments from Mongolia and Japan (L2/15-009).

Recommendations: We recommend the UTC review this document at their leisure. The topics raised in the response by Pandey are expected to be discussed at a meeting with the Mongolians and Japanese representatives in Tokyo, Pandey, and WG2 members before the next WG2 meeting in Matsue, Japan.
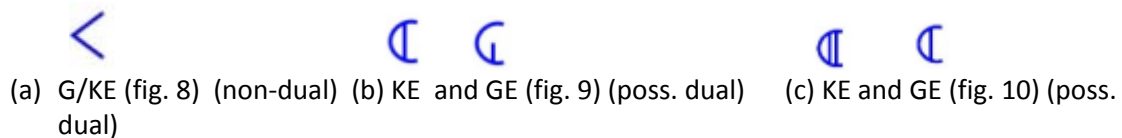
EUROPE
10. North Eastern Iberian
Document:
L2/15-120 Northeast Iberian (revision of earlier L2/15-012) – Joan Ferrer, et al., LITTERA group, University of Barcelona

Discussion: We reviewed this proposal, which revised the earlier document (L2/15-012), based on feedback from the February UTC meeting.

The script ad hoc had the following comments:
- Based on the text in the proposal, the script name "IBERIAN" (i.e., without "Northeast") seems warranted
- In the Introduction, mention that "Levantine Iberian" has been used when referring to this script
- What is the justification for the order of the characters a1, a2 (given in the listing on pp. 9-10)?
- On page 9, create a table with a transcription scheme for the character names (in ASCII), avoiding numbers in the names
- Because a "/" is not allowed in character names, use FRACTION ONE QUARTER or FRACTION ONE HALF
- Provide more examples of the numbers, with a clear description showing how the numbers would be represented (with code points)
- For the non-dual forms, how were the representative glyphs selected?
- In the examples below, can the authors explain how the values were arrived at, and why KE in (c) below gets two bars, but GE one (whereas in figure 9, KE has only one bar)?

(a) G/KE (fig. 8) (non-dual) (b) KE and GE (fig. 9) (poss. dual)   (c) KE and GE (fig. 10) (poss. dual)

Recommendations: We recommend the UTC review this document and forward any feedback (along with the comments above) to the authors.

11. Southern Palaeohispanic
Document: L2/15-119 Southern Palaeohispanic - Joan Ferrer, et al., LITTERA group, University of Barcelona

Discussion: We reviewed this preliminary proposal for Southern Palaeohispanic.

In our opinion, a strong case needs to be made for encoding Southern Palaeohispanic separately. The situation is reminiscent of the many alphabets in ancient Italy, which were unified as "Old Italic" (cf. the chart at: http://www.unicode.org/L2/L2000/00140-old-italic.pdf).

In addition, can the authors explain clearly what the reference letters and numbers indicate in the text? Are they catalog numbers? For example, on page 1 "J.25.1" in "(Untermann 1997 J.25.1)" and "H.3.1." in the caption for figure 10, "Silver vase from Santiesteban del Puerto (H.3.1)".

Recommendations: We recommend the UTC review this document and forward any feedback (along with the comments above) to the authors.

**AFRICA**
12. Mandombe

Document: [L2/15-118](#) Preliminary Proposal for Encoding the Mandombe Script in the SMP – Rovenchak et al.

Discussion: We reviewed this preliminary proposal.

The ad hoc discussion raised the following points:
- Since the topic of the copyright has come up in the past, the authors should address it in the proposal.
- We noted that 20 glyphs for the calendar characters (which begin with KIMBANGUMUNA) are missing.
- Can the authors provide information on the user community actively using the script? (Wikipedia says the the Kimbanguist church has 5.5 million believers.)
- The proposal should include details on how the script works, so it can be implemented on a computer. The type of information required is along the lines of that contained on the Wikipedia page, which describes the four basic consonant shapes, such as:

Consonant	Family 1	Family 2	Family 3	Family 4
ꚅ Group 1	ꚅ na	ꚅ va	ꚅ sa	ꚅ ta

Recommendations: We recommend the UTC review this proposal and send the authors feedback.

13. Medefaidrin
Document: [L2/15-117](#) Medefaidrin – Rovenchak et al.

Discussion: We reviewed the preliminary proposal for the Medefaidrin script, a script specifically created for a special-purpose language used by the Obɛri Ɔkaimɛ religious community in the Ibibio community of Nigeria. The script is part of an ongoing digitization project, in an effort to document this extremely endangered language and its script. The language has attracted the attention of linguists because it is a rare example of a language whose entire life-cycle can be studied: it has gone from birth to near extinction within 80 years.

The script ad hoc discussion raised the following points:
- Background on the script: How many people are in the user community and how many people use it? Is it actively taught?
- The number of columns should be reduced from 7 to 6
- The characters for ten to nineteen should be renamed NUMBER (and not DIGIT)
- The PERIOD character should be renamed FULL STOP
- The characters ONE COMBINED, TWO COMBINED, THREE COMBINED should be renamed to ONE ALTERNATE FORM, TWO ALTERNATE FORM, etc.
- The analysis of the numbers needs to be more complete:
  - Discuss 20, which has a left-tic on it but is not discussed in the proposal
  - Show how 40, 41, 42, 90, and 100 are represented

Recommendations: We recommend the UTC review this proposal and send any feedback (along with the comments above) to the authors.

14. Egyptian
Documents: [L2/15-069](#) Egyptian Hieroglyphs in Unicode plain text, a note on a suggested approach – Richmond
[L2/15-123](#) Proposal to encode three control characters for Egyptian Hieroglyph – Richmond

Discussion: We reviewed these two documents: the background document ([L2/15-069](#)) and the proposal ([L2/15-123](#)). The proposal requests three Egyptian Hieroglyph Sign format characters be added. These characters will cluster the signs into Egyptian quadrats, which are the rectangular clusters of hieroglyphs arranged in an aesthetically pleasing way. The three characters proposed are equivalent to the MdC control codes: "*" (juxtaposer), ":" (subordinator) and "&" (ligator, proposed with a "+" glyph in L2/15-123).

The script ad hoc had a spirited discussion and touched on the following points:
- The original Egyptian repertoire was encoded as a set of text elements, with the understanding that layout of characters was not plain text encoding, and markup (MdC) would be used to represent the arrangement of signs (cf. mathematicians who rely on markup).
- Current mechanisms are not working for Egyptologists, meaning that the user community may create their own, non-standard approaches.
- Do we need more control functions than what we already have? What additional needs are required that ZWJ and ZWNJ don't provide? (The answer: According to users, three virama characters are deemed necessary for clustering.)
- Shorthand format controls were encoded for Duployan, so can format controls be added for Egyptian Hieroglyphs? Are Egyptologists willing to wait for 5 years for such characters to be approved and implemented?
- If current advances in rendering have meant that clustering (with such characters) is within reach of plain text mechanisms, then the UTC should re-visit the addition of format controls for Egyptian hieroglyphs.

Recommendations: We recommend the UTC discuss this, since it involves a change to the encoding model for Egyptian.
Note: The topic is very timely, given that there are ongoing discussions about Mayan hieroglyphs, which has a more complicated layout than Egyptian hieroglyphs.

MIDDLE EAST

15. Arabic
Document: [L2/15-087](#) Proposal to encode one Arabic honorifics - Lateef Sagar Shaikh

Discussion: We reviewed this request for one Arabic honorific, similar to the proposal for seventeen honorifics by Roozbeh Pournader ([L2/14-147](#)).

The proposal requests one character, ARABIC LIGATURE KARAM ALLAHU WAJHULKARIM. However, the evidence presented does not show the character as a combined ligature. Also, no evidence is provided of the character in a font.

Recommendations: We recommend the UTC respond to the author, and request he provide more attestation of the ligature in plain text and evidence of the character in a font.

16. Syriac
a) Combining marks

Document: L2/15-081 Syriac appears to be missing various combining characters – Pournader

Discussion: We quickly reviewed this document, which provides links to modern Syriac documents. The documents contain diacritical marks whose representation may not be specified in the Syriac block introduction of the Unicode Standard.

(It should also be noted that diacritics in Syriac in general should be examined thoroughly, to ensure they can be represented by Unicode characters.)

Recommendations: We recommend the UTC note this document, and encourage interested parties to conduct further research to see if additional text should be added to the standard, and/or whether new characters need to be proposed.

b) Letters for Garshuni Malayalam
Document: L2/15-088 Preliminary Proposal to Encode Syriac Letters for Garshuni Malayalam – Pandey

Discussion: We reviewed this preliminary proposal, which proposes 11 letters for Garshuni Malayalam, a way of writing the Malayalam language in a variant of Syriac script. Garshuni Malayalam was popular amongst Saint Thomas Christians of Kerala, India. The proposal has had input from experts.

In review of the evidence presented, it appears that only one new column for "Syriac Extensions" is needed, although this proposal shows 2 columns (U+0860..U+087F). Also, the joining classes of the 11 proposed characters need be reviewed.

Recommendations: We recommend the UTC review this document and solicit additional feedback from their members, forwarding any comments to the author.

NUMBER SYSTEMS

17. Ottoman Siyaq Numbers
Document: L2/15-072 Proposal to Encode Ottoman Siyaq Numbers – Pandey

Discussion: We quickly reviewed the proposal for Ottoman Siyaq Numbers, which was initially discussed in 2007 and put forward as a separate block in 2009. The characters are well-attested through secondary sources, although it has not been possible to locate the original documents. As the final proposal was not available to us at the time of the ad hoc meeting, we could not give it a closer review, but look forward to the final version at the UTC.

Recommendations: We recommend the UTC carefully review the final proposal, and discuss whether all four Siyaq blocks be given the Arabic script property, or whether each block should receive its own script property.

18. Diwani Siyaq Numbers
Document: L2/15-066 Proposal to Encode Diwani Siyaq Numbers – Pandey

Discussion: We quickly reviewed the proposal for Diwani Siyaq Numbers, which was first discussed in 2007. The characters are well-attested, but not through original materials, only secondary sources. We look forward to the final proposal at the UTC.

Recommendations: We recommend the UTC carefully review the final proposal, and discuss whether all four Siyaq blocks be given the Arabic script property, or whether each block should receive its own script property.

19. Indic Siyaq Numbers
Document: L2/15-121 Proposal to Encode Indic Siyaq Numbers - Pandey

Discussion: We quickly reviewed the proposal for Indic Siyaq Numbers, which was earlier called "Raqm". Like the other set of Siyaq numbers, it was first discussed in 2007. We look forward to the final proposal at the UTC.

Recommendations: We recommend the UTC carefully review the final proposal. The UTC should also discuss whether all four Siyaq blocks be given the Arabic script property, or whether each block should receive its own script property.

(NOTE: L2/15-122 **Proposal to Encode Persian Siyaq Numbers** was not reviewed by the script ad hoc committee as it had not yet been submitted for posting**.)**

20. Siyaq Number Mark
Document: L2/15-074 Siyaq Number Mark - Pandey

Discussion: We reviewed this request for one character, U+ 061D ARABIC SIYAQ NUMBER MARK, which appears over numbers and indicates that the numbers belong together as a set.

The script ad hoc recommended the proposal include wording describing what characters it appears on top of, such as: "This character appears over any sequence of Siyaq numbers, from any Siyaq block, and full stop". Additionally, can the author specify the maximum number of characters over which it can extend? Does it ever go across a word break? (If so, provide evidence.)

Recommendations: We recommend the UTC discuss this proposal and approve U+ 061D ARABIC SIYAQ NUMBER MARK, after modifications based on the comments above and discussion in the UTC.

NOTATIONAL SYSTEMS
21. Pitman Shorthand
Document: L2/15-116 Pitman (revision of L2/14-254) Rajaram

Discussion: We took a cursory look at this proposal, which has been revised based on comments from the February UTC. The proposal is currently being reviewed by other Pitman users.

Recommendations: We recommend the UTC review this proposal, and send feedback to the author.

SYMBOLS
22. Various Proposals for Symbols

a) L2/15-093 Proposal to Encode 'Shiva Linga' Symbols - Pandey

b) L2/15-099 Proposal to Encode the 'Fravahar' Symbol - Pandey

c) L2/15-097 Proposal to Encode Symbols of Jainism - Pandey

d) L2/15-100 Proposal to Encode Winged Sun Symbol - Pandey

e) L2/15-105 Proposal to Encode the 'Atashdan' Symbol - Pandey

f) L2/15-111 Proposal to Encode the 'Parsapen' Symbol - Pandey

Discussion: We briefly reviewed the above documents, which request various religious symbols be encoded. The symbols are identified with Hinduism (Shiva Linga in L2/15-093), Jainism (4 symbols in L2/15-097), Zoroastrianism (Fravahar in L2/15-099, Atashdan in L2/15-105), and for the religion of the indigenous Gond community (Parsapen in L2/15-111).

As no set of guidelines has been put forward by the emoji subcommittee or the UTC regarding the encoding of religious symbols, it wasn't clear to us whether the script ad hoc should deal with them, or the emoji subcommittee. However, since this set is similar to emoji (such as the MENORAH WITH NINE BRANCHES), we recommend they be discussed within the emoji subcommittee. For those symbols that apply to maps, we recommend they be folded into the work being done by Shervin Afshar (see Action Item 142-A45 and Appendix 2 of L2/14-235R3).

Recommendation: We recommend these proposals go to the emoji subcommittee, and characters appearing on maps (i.e., those characters appearing on maps made of paper and which have a history), be relayed to Shervin Afshar for future proposals (see Action Item 142-A45).

23. Group Mark
Document: L2/15-083 Proposal for addition of Group Mark symbol - Ken Shirriff

Discussion: We reviewed this proposal for the Group Mark symbol, which was originally used as a separator character for I/O operations in the 1950s, and then became a part of computer character sets in the 1960s and 1970s. The character has appeared in manuals up to the present day. The request for encoding this symbol is based on its usage in running text (and not as a control character).

The proposal appears to be well-formed, with extensive examples and discussion. The location of the character needs to be changed, however, since the proposed code point U+23FB is currently taken by

the POWER SYMBOL. (The properties on page 6 list both U+23FA and U+23FB as the code point; presumably U+23FA, at the top of the page is a typo.)

Recommendations: We recommend the UTC approve the GROUP MARK, but request the author update the proposal with the code point U+2BD2.

24 Hebrew Nomina Sacra
Document: L2/15-092 Typographic Concerns and the Hebrew Nomina Sacra - Shoulson

Discussion: We reviewed this preliminary document, which gives a selection of examples showing how the Tetragrammaton appears in Hebrew, but whose representations cannot be handled by current Unicode characters.

Based on the document, it appears that the encoding of two characters would make it possible to represent many of the examples. The characters are (with the following approximate glyphs):

(1)   (2)

Recommendations: We recommend the UTC discuss this proposal and respond to the author, based on the comments above and UTC discussion.

Previous recommendations (carried over for script and character proposals not yet discussed in the UTC)

--- CARRIED OVER ---
EAST AND CENTRAL ASIA
12. Small Seal Script
L2/14-242 Proposal to encode Small Seal Script – TCA and China

Discussion: We reviewed this proposal, which proposes 799 characters out of a projected 10,516. In our opinion, the proposal is still far from mature, and would benefit from coordinating work with experts in the U.S. and Japan in order to formalize mapping data, which is needed to evaluate a final proposal. The proposal should also provide demonstrated need for including the script in the international standard.

Recommendations: We recommend the UTC members review this proposal and consider sending the authors the comments above.

13. Naxi Dongba
L2/14-241 Supplement on Proposal for Encoding Naxi Dongba Pictograph Script (L2/11-178) - China
L2/14-245 Feedback on Naxi Dongba Supplement document - Anderson

Discussion: We reviewed the "Supplement" document, which answered questions posed at the June 2011 WG2 meeting in Helsinki, Finland (see Naxi Dongba Ad Hoc report, L2/11-244). Specifically, the

authors in the "Supplement" confirmed that the encoding is for modern use, not traditional use of the characters, and that alphabetical ordering is preferred.

The "Feedback" document posed additional questions and made suggestions. During WG2 discussion, the Naxi Dongba proposal authors stated the script is both a logography and syllabary, and the variation shown in some glyphs is due to regional differences, but only one glyph per character is warranted in the encoding. They agreed to revise the proposal and provide information on the proposed characters, with glyphs, Romanized transcription, Chinese gloss (and English translation) and references.

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

14. Shuishu
L2/14-243 Proposal for encoding Shuishu – China

Discussion: We reviewed this proposal, which is still at an early stage. In our view, it is not yet clear that Shuishu is an encodable writing system. In order to move forward, we recommend the authors prepare and publish a standard sign list for Shuishu, which can then be circulated for review by other scholars and gain scholarly support. The next version of the proposal should also provide a rationale for the digital representation of their sign list, answering the question why these shapes should be put into an international character encoding standard.

Recommendations: We recommend the UTC members review this proposal and send comments to the authors. The UTC may want to relay the suggestions to the authors above, regarding recommended next steps.

15. Khitan Large Script
L2/14-234 Proposal on Encoding Khitan Large Script – China
L2/14-233 Preliminary Review of Proposal on Encoding Khitan Large Script – West
L2/14-246 Ad hoc reports for Tangut and Khitan Large Script – Anderson

Discussion: We reviewed these documents. As noted in L2/14-233, the Khitan Large Script is largely undeciphered without any character list or recent dictionaries, vocabulary lists, or secondary linguistic materials, so the current proposal should be viewed as preliminary.

Also as mentioned in L2/14-233, the script appears to have a significant percentage of characters (18%) that are either Han borrowings or identical in shape to already encoded CJK ideographs. A revised proposal should discuss the pros and cons of unifying those Khitan Large Script characters with CJK characters already encoded: what are the costs/benefits to unification? Because Khitan Large Script is an historical script, the security risk would not arise if Khitan Large Script used CJK characters, only if it encoded a large set of identical CJK characters.

Additionally, we suggest the proposal also create a "Uni-Khitan" database (or spreadsheet) to document sources.

Recommendations: We recommend the UTC members discuss these documents.

16. Ranjana

L2/09-192 Preliminary proposal for encoding the Rañjana script in the SMP (WG2 N3649)
L2/14-221 Comparison between Ranjana Proposals - Anderson
L2/13-243 Proposal to Encode Ranjana Script - Manandhar
L2/14-253 Recommendations to UTC from Script Meeting in Nepal - Anderson

Discussion: We discussed these documents. Since decisions on the repertoire and encoding model for Ranjana depend upon those for "Nepaalalipi", discussion on Ranjana was limited. It was noted that a future Ranjana proposal should also discuss the unification with Wartu and Lanydza, and should provide details on any specific characters and behaviors of the script in Tibet and other locations outside Nepal.

Recommendations: We recommend the UTC review the document, but postpone discussion until after the "Nepaalalipi" encoding is resolved.

17. Bhujinmola

L2/14-253 Recommendations to UTC from Script Meeting in Nepal
L2/14-283 Introducing the Bhujinmol Script - Pandey

Discussion: We briefly discussed the section in the "Recommendations" on Bhujinmola. Bhujinmola has a characteristic wavy headline (see examples in "Roadmapping the Scripts of Nepal" L2/09-325). The question on whether Bhujinmola represents a stylistic variation of "Nepaalalipi" or should be separately encoded needs to be discussed in a separate document, with examples of how vowels and consonants join differently from "Nepaalalipi" and other rendering issues.

Recommendations: We recommend the UTC review the document, but wait for further research to support separately encoding Bhujinmola. (Note: The script ad hoc did not yet review L2/14-283 Introducing the Bhujinmol Script by Pandey.)