

To: UTC

L2/12-378

From: Deborah Anderson, Rick McGowan, and Ken Whistler

Title: Review of Indic-related documents and Recommendations to the UTC

Date: 5 November 2012

Note: This review does not include scripts that were approved by WG2 and are contained in the Consent Docket (L2/12-344).

I. NEWAR/NEPAALALIPI

Documents:

L2/12-003 Proposal to encode the Newar Script (revised) – Anshuman Pandey

L2/12-120 Proposal for the Nepaalalipi script in the UCS - Dev Dass Manandhar, et al.

L2/12-336 Letter in support of N4184 [L2/12-003] and encoding the Newar script (WG2 N4372) [L2/12-120] – Iain Sinclair

L2/12-349 Proposal for the Nepaalalipi script – Dev Dass Manandhar, et al.

Discussion: We reviewed the various documents.

The latest proposal from Dev Dass Manandhar (L2/12-349) now includes the encoding of long I and U atomically, but also four different formats of each six vowels (see page 5 in L2/12-349), which results in the addition of characters beyond those proposed in his earlier document (L2-12-120).

In evaluating the proposals, we direct the UTC to review the analysis by Iain Sinclair in evaluating the two competing proposals (though note that Dr. Sinclair was only able to see the earlier proposal of Dev Dass Manandhar, L2/12-120, not his more recent version).

Recommendation: As there will be a meeting between the authors of the proposals for this script this month (November), we recommend the UTC hold off on approving any proposal, but encourage the authors to work out their differences. We also suggest the UTC discuss and consider sending comments on the new vowel characters as proposed in Dev Dass Manandhar's document L2/12-349.

II. NEPAL HIMALAYISH

Documents:

L2/12-365 Proposal to Encode Nepal Himalayish Scripts – Pat Hall

L2/12-336 Letter in support of N4184 [L2/12-003] and encoding the Newar script (WG2 N4372) [L2/12-120] – Iain Sinclair

Discussion: We reviewed these documents. The proposal for Nepal Himalayish puts forward a Devanagari-based encoding, relying on fonts to represent the scripts of the various Tibeto-Burman languages of Nepal. Such an approach, which relies on language family classification within a geographical and political boundary, should, we feel, instead look closely at the behavior of each script, such as the conjunct behavior of consonants, rendering behavior in general, and other features.

The letter from Iain Sinclair includes a few critical remarks on Himalayish (see pages 2-3).

Recommendation: We recommend the UTC respond to the author, underlining Unicode’s approach to encoding (i.e., that Unicode encodes *scripts*, not languages - or sounds- and that Unicode does not define orthographies for languages, but characters for scripts).

III. ASSAMESE

Documents:

L2/12-339 Two Technical problems Concerning Assamese - Jyotiprakash Tamuli

L2/12-350 Proposal for minor, non-character, additions to UCS addressing concerns from Assamese - Martin Hosken

Discussion: We reviewed L2/12-339, which identifies concerns of the Assamese community, and L2/12-350, which suggests proposed solutions to those concerns.

The proposed solutions in L2/12-350 include adding:

- a paragraph in the block header (in the names list), describing the close relationship between Bengali and Assamese, with input from interested parties
- a named sequence for Assamese ksha:
BENGALI LETTER ASSAMESE KSSA;0995 09CD 09B7
- two name aliases:
ASSAMESE RA for 09F0 BENGALI LETTER RA WITH MIDDLE DIAGONAL
ASSAMESE WA for 09F1 BENGALI LETTER RA WITH LOWER DIAGONAL
- documentation of the appropriate rendering of two conjuncts specific to Assamese:
09F0 BENGALI LETTER RA WITH MIDDLE DIAGONAL + 09C1 BENGALI VOWEL SIGN U
09F0 BENGALI LETTER RA WITH MIDDLE DIAGONAL + 09C2 BENGALI VOWEL SIGN UU

Recommendation: We recommend the UTC discuss the issues raised in these documents, and remand the relevant items to the Editorial Committee.

IV. SINHALA

Document: L2/12-358 Issues with Sinhala symbols Rakaaraansaya and Yansaya – Gihan Dihas

Discussion: We reviewed L2/12-358. The issues confronting Sinhala are laid out in this document.

Recommendation: We recommend the UTC discuss this document in plenary.

V. GRANTHA

Documents:

L2/12-039 Proposal to add two letters to the Grantha repertoire – Everson

L2/12-348 Comments on Proposal L2/12-039, "Proposal to add two letters to the Grantha repertoire" - S. Palaniappan

L2/12-353 Letters LLLA and NNNA in Grantha block - Reference documents from experts and users - Naga Ganesan

Discussion: We reviewed the original proposal for two additional Grantha characters, NNNA and LLLA (L2/12-039), L2/12-348, which counsels against encoding Grantha NNNA and LLLA, and L2/12-353, which recommends encoding them.

Recommendation: We recommend the UTC await feedback from other interested parties. Note that in L2/12-180, dated 5 May 2012, INFITT had expressed an interest in commenting on L2/12-039.

SCRIPT TOPICS CARRIED OVER FROM AUGUST 2012 UTC MEETING

VI. TAMIL

Documents:

L2/12-231 Proposal to encode Tamil fractions and symbols - Shriramana Sharma

L2/12-232 Regarding my Tamil fractions/symbols proposal - Shriramana Sharma

Discussion: We only quickly reviewed the proposal itself, L2/12-231, which is long (66 pages) and complicated.

In §3 of L2/12-232, the author requests the UTC take no action on the Tamil fractions and symbols proposal (L2/12-231), as he is soliciting feedback from relevant parties in India and Sri Lanka.

Recommendation:

In view of §3 in L2/12-232, we recommend the UTC members review the main proposal, L2/12-231, and if they have any expertise or knowledge of details of Tamil fractions and symbols, they should relay feedback to the document's author.

VII. SHARADA

Document: L2/12-224 Annotation for 111C4 SHARADA OM - Shriramana Sharma

Discussion: We reviewed this proposal.

Recommendation: We recommend the UTC note this short document and suggest it go to the Editorial Committee as feedback on the names list.

SCRIPT TOPICS CARRIED OVER FROM MAY 2012 UTC MEETING

VIII. GENERIC INDIC COMMENTS

Document: L2/12-106 Request for editorial updates to various Indic scripts, §1 - Shriramana Sharma

a. L2/12-147 Review of Indic-related L2 documents and Recommendations to the UTC

Discussion: We reviewed §1 of L2/12-106.

Recommendation: While the proposed text changes to the text seem reasonable, we recommend UTC members, particularly those with rendering engines for Devanagari, review this request and verify the proposed change in the text of Chapter 9 is acceptable.

b. L2/12-184 from GOI: The proposal is being referred to experts to give their views.

IX. TAMIL

Document: L2/12-106 Request for editorial updates to various Indic scripts, §2 Tamil Debit and Credit - Shriramana Sharma

a. L2/12-147 Review of Indic-related L2 documents and Recommendations to the UTC

Discussion: We reviewed §2 of L2/12-106.

Recommendation: We recommend the glyph for U+0BF6 TAMIL DEBIT SIGN be changed, based upon this document and personal email feedback from Uma Umamaheswaran, who was an author of the original proposal (with INFITT), L2/01-375R. Additional feedback from INFITT would also be welcome. For U+0BF7 TAMIL CREDIT SIGN, we recommend additional feedback be sought from UTC members and other interested parties.

b. L2/12-180 from INFITT:

[...] Since the original Tamil Debit Sign glyph is based on the Government of Tamil Nadu 8-bit standards for Tamil and the proposal that INFITT co-authored with Dr. Uma Umamaheswaran, INFITT has an interest in reviewing Mr. Shriramana Sharma's proposal (L2/12-106). INFITT's WG02 members Mr. P. Chellappan (who provided the glyphs in the INFITT-Uma proposal) and Dr. Ramakrishnan are part of the user community that uses the debit sign in their accounting. They insist that the shape in the INFITT-Uma proposal is what they learned to use. Still, INFITT would like to review this further and offer up additional comments once we have done more research and have had an opportunity to consult additional experts in the user community. We expect to submit our comments before the August UTC meeting.

c. L2/12-184 from GOI:

The proposal is being referred to experts to give their views. Hence we request the UTC to defer the decision on the above proposals till next UTC meeting

X. MALAYALAM

Document: L2/12-106 Request for editorial updates to various Indic scripts, §3 Malayalam - Shriramana Sharma

a. L2/12-147 Review of Indic-related L2 documents and Recommendations to the UTC

Discussion: We reviewed §3 of L2/12-106.

Recommendation: We recommend the Malayalam block intro text be modified to acknowledge the glyphic form of the dot reph as a small vertical stroke, besides the dot shape. Before making changes to the text regarding C2 conjoining forms, we suggest the UTC solicit feedback from members with Malayalam rendering engines.

L2/12-184 from GOI:

The proposal is being referred to experts to give their views. Hence we request the UTC to defer the decision on the above proposals till next UTC meeting

XI. ORIYA

Document: L2/12-106 Request for editorial updates to various Indic scripts, §4 Oriya - Shriramana Sharma

a. L2/12-147 Review of Indic-related L2 documents and Recommendations to the UTC

Discussion: We reviewed §4 of L2/12-106.

Recommendation: Because we have no expertise on current rendering engines, we defer to the UTC on the proposed wording change, specifically to get feedback from those members with working Oriya implementations.

b. L2/12-184 from GOI:

The proposal is being referred to experts to give their views. Hence we request the UTC to defer the decision on the above proposals till next UTC meeting

FOR UTC MEMBERS' INFORMATION ONLY

XII. KANNADA

Document: L2/12-203 Supporting Tulu language written in the Kannada script - Shriramana Sharma

Discussion: We reviewed this document, which is primarily a “FYI” document to UTC members regarding a potential revision to the orthography for the Tulu language in the Kannada script.

Recommendation: We recommend the UTC take no action, but members with Kannada rendering engines should be on notice that the orthography for the Tulu language in the Kannada script is complicated, and the representation for certain vowels may change.

XIII. GONDI

Document: L2/12-235 Revised Preliminary Proposal to Encode the Gondi Script - Anshuman Pandey

Discussion: We reviewed this document, which is a revision of an earlier preliminary proposal. We note the script is already on the Roadmap.

Recommendation: We recommend UTC members review this proposal and send the author any feedback.

XIV. ROHINGYA

Document: L2/12-214 Preliminary Proposal to Encode the Rohingya Script - Anshuman Pandey
2012-06-25

Discussion: We reviewed this document, which is preliminary only. Note that the script proposal includes a TATWEEL. Additional information on the extent of its use should be included in the next version of the proposal, along with a discussion of whether a specific TATWEEL should be separately encoded.

Recommendation: We recommend UTC members review this proposal and send the author feedback.

XV. COORGI-COX

Document: L2/12-217 **Introducing the Coorgi-Cox Alphabet - Anshuman Pandey**

Discussion: We reviewed this preliminary proposal. The script has not yet been put on the Roadmap, which seems appropriate since the extent of its use is not yet verified.

Recommendation: We recommend UTC members review this proposal and send the author feedback.