

To: South Asian Subcommittee and UTC

L2/11-298

From: Deborah Anderson, Rick McGowan, and Ken Whistler

Title: Review of Indic-related L2 documents and Recommendations to the UTC

Date: 25 July 2011

I. GRANTHA

Documents: L2/11-182, L2/10-186, L2/11-200, L2/11-284

Discussion:

We reviewed the current L2 documents relating to the encoding of Grantha, which have appeared in the document registry since the last UTC.

These include:

- L2/11-284 Revised proposal from Government of India (M. Jain, Gov't of India)
- L2/11-200 Grantha Compromise Chart (McGowan/Anderson)
- L2/11-186 Dravidian Letters in Grantha (N.Ganesan)
- L2/11-182 Letter to Mark Davis (N. Ravi Shanker, Gov't of India)

L2/11-284 is the revised proposal from the Government of India, which has taken into account the Government of Tamil Nadu's feedback. As such, it provides the official Government of India response which was mentioned in L2/11-182, a letter from N. Ravi Shanker. (The letter from N Ravi Shanker, dated 9 May 2011, had requested Unicode Consortium should delay a decision until Competent Authority in the Government of India had approved the recommendations on Grantha.)

L2/11-284 has been compared to the Grantha Compromise Chart in L2/11-200. The two documents are identical in their codepoints, glyphs, and names, with the exception of one character, U+1135D GRANTHA SIGN PLUTA, which appears only in the Compromise Chart.

The letter from N. Ganesan (L2/11-186) on the Dravidian Letters in Grantha was noted.

Recommendation: We recommend the UTC approve Grantha, based on the proposal from the Government of India. We recommend the additional character (GRANTHA SIGN PLUTA), which the Government of India had earlier approved (L2/10-409), be dealt with as a ballot comment in WG2, which the Government of India can also note in their ballot comments.

II. OTHER SOUTH ASIAN TOPICS

1. Vedic

Document: L2/11-267 Proposal to encode svara markers for the Jaiminiya Archika (Shriramana Sharma), L2/11-289 Comments on Public Review Issues

Discussion: We reviewed L2/11-267 and the feedback from John Cowan, contained in the Comments on Public Review document L2/11-289.

Recommendation: We recommend the two characters, proposed U+1CF8 VEDIC TONE RING ABOVE and U+1CF9 VEDIC TONE DOUBLE RING ABOVE in L2/11-267 be approved.

We consider the annotations for U+0951 and U+1CD2 (“also used in Jaiminiya Sama Archika”) to be reasonable input for the Editorial Committee.

2. Mahajani

Document: L2/11-274 Proposal to Encode the Mahajani Script (Anshuman Pandey)

Discussion: We reviewed L2/11-274 and find the proposal to be mature.

Recommendation: We recommend the UTC approve this script.

3. Bhaiksuki

Document: L2/11-259 Preliminary Proposal to Encode the Bhaiksuki Script (Anshuman Pandey)

Discussion: The proposal is missing glyphs for the digits “0” and “3” at U+14450 and U+14453. Since these two numbers are not known to be missing in any of the scripts in this area, we suggest the code chart show a box (or some other glyph) to reflect that no attestation has yet been found. The name for U+1446C should be changed from “BHAIKSUKI LETTER-NUMERAL HUNDREDS” to “BHAIKSUKI LETTER-NUMERAL HUNDRED.”

Recommendation: We recommend others in the UTC review the proposal and forward their comments to the proposal author.

4. OM

a. Oriya

Document: L2/11-258 Proposal to encode 0B50 ORIYA OM (Shriramana Sharma)

Discussion: We reviewed this proposal, which provides evidence of ligated forms.

Recommendation: We recommend the UTC approve this character.

b. Bengali

Document: L2/11-275 Bengali Om character (Barun Kumar Sahu)

Discussion: We reviewed the proposal.

Recommendation: We recommend the UTC accept this character.

5. Jenticha

Document: L2/11-218 Proposal to Encode the Jenticha Script (Anshuman Pandey)

Discussion: In section 1, “Introduction,” the first sentence should be revised, with the word “preliminary” removed.

In section 3.5, “Vowel Length,” the colon-like character is mentioned as being used in the modern script, but not distinguished in the older, alphabetic script. It appears in the children’s book in figure 11 in the Jenticha script, corresponding to a Devanagari visarga. What is its use here in both scripts? The colon also appears with parentheses in the Jenticha text in figure 12, but does not appear in the Devanagari. How is it being used here? More information on this character is needed.

In section 3.13, “Digits,” the proposal should call out the early shapes of the digits ‘0’, ‘2’, and ‘3’ are different from the glyphs in the code chart. (The early shapes of these digits are exemplified in Table 4.)

Figure 8 on page 12 needs to be re-done, so it displays better.

Recommendation: We recommend the UTC approve this script, with the proviso that a revised version of the proposal, which includes the edits and additional documentation requested above, be incorporated, before being sent to WG2.

6. Pau Cin Hau

Document: L2/11-287 Proposal to Change the Names for Some Pau Cin Hau Characters (Anshuman Pandey)

Discussion: We have reviewed the request to change the names for six Pau Cin Hau characters:

[Amendment 1.2 charts in N4107]

11AC5 Y

11ACA X

11AE7 PAU CIN HAU GLOTTAL STOP VARIANT

11AEA PAU CIN HAU GLOTTAL STOP VARIANT FINAL

11AF0 PAU CIN HAU GLOTTAL STOP OTHER

11AF1 PAU CIN HAU MID-LEVEL TONE LONG

[Changes as shown in proposal L2/11-104R]

Z

KH

SANDHI GLOTTAL STOP

SANDHI GLOTTAL STOP FINAL

GLOTTAL STOP VARIANT

MID-LEVEL TONE LONG FINAL

Recommendation: We recommend the UTC accept the six Pau Cin Hau name changes and make ballot comments accordingly.

7. Siyaq

a. Persian Siyaq Numbers

Document: L2/11-272 Preliminary Proposal to Encode Persian Siyaq Numbers in the UCS (Anshuman Pandey)

Discussion: We briefly reviewed the proposal for Persian Siyaq, which is perhaps the best attested of the set of four Siyaq “scripts.”

Recommendation: We recommend the UTC review this proposal, and request the Roadmap be updated to match the code point range in the proposal (1EC00-1EC6F).

b. Diwani Siyaq Numbers, Ottoman Siyaq Numbers, Indic Siyaq Numbers

Documents: L2/11-269 Preliminary Proposal to Encode Diwani Siyaq Numbers in the UCS (A. Pandey), L2/11-271 Preliminary Proposal to Encode Ottoman Siyaq Numbers in the UCS (A. Pandey), L2/11-270 Preliminary Proposal to Encode Indic Siyaq Numbers in the UCS (A. Pandey)

Discussion: We briefly reviewed the proposals.

Recommendation: We recommend the code points in these three proposals be changed as follows: move Indic Siyaq to 1EC70-1ECBF, Diwani Siyaq to 1ECC0-1ECFF, and Ottoman Siyaq to 1ED00-1ED3F. We suggest the proposals with their revised code points be re-submitted for consideration at the November UTC meeting. We also suggest UTC members review the proposals and forward their feedback to the proposal author.