To: UTC  
From: Deborah Anderson, Rick McGowan, and Ken Whistler  
Title: Review of Indic-related L2 documents and Recommendations to the UTC  
Date: 25 April 2012

I. INDIC
Documents:
L2/12-106 Request for editorial updates to various Indic scripts, §1 Generic Indic - Shriramana Sharma
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.

II. TAMIL
Document: L2/12-106 Request for editorial updates to various Indic scripts, §2 Tamil - Shriramana Sharma
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.

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

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

V. BRAHMI
Document: L2/12-106 Request for editorial updates to various Indic scripts, §5 Brahmi - Shriramana Sharma
Discussion: We reviewed §5 of L2/12-106.
Recommendation: We recommend the UTC approve this glyph change.

VI. “Nepālalipi” / “Newar”

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

Discussion: We reviewed the two documents. The proposals largely share the same character repertoire (cf. charts on p. 20 of both proposals), but there are three significant differences in the encoding model, specifically how to represent (1) long vowels I and U, (2) diphthongs AI and AU, and (3) the six “resonant breathy” consonants (NGHA, NJHA, NHA, MHA, RHA, and LHA).

(1) The long vowels I and U (both dependent and independent) are encoded as separate characters in L2/12-003, but are handled as ligated forms of vowel + visarga (“liphuti”) in L2/12-120 (see top of page 10, and Tables 13 and 14).

Collation behavior may help distinguish between the two alternatives. If the user community desires the typical Brahmī-based script ordering of long I and U: i < ii < u < uu, then the simplest way to achieve that is by atomic encoding of the long I and long U. If, on the other hand, primary ordering of long I and long U is not that important to the user community, then the default ordering resulting from encoding these long vowels as sequences <i, visarga>, <u, visarga> might be sufficient. However, this may also depend on the encoding of the dependent (matra) vowels. If the dependent long -I and long -U are separately encoded, then it is inadvisable to encode the independent long I and long U vowels as sequences, merely because of their glyphic appearances.

(2) The diphthongs AI and AU are encoded as independent vowels in L2/12-003, but appear to be handled in L2/12-120 as ligated forms of A with dependent forms of AI or AU (see section 7.3 and Table 11, p. 14). As for the long vowels I and U, if collation is not important, then the use of a sequence for AI and AU might be acceptable. However, we may have misunderstood the encoding intent in L2/12-120 for these diphthongs, as the text in §7.3 and §10 was found to be confusing.

We note that the typical Brahmī order is: a < aa < i < ii < u < uu < e < ai < o < au (and is reflected in L2/12-003). Document L2/12-120 seems to suggest: a < aa < i < ai < u < au < e < o. However, if long I and long U are not atomically encoded, then the order would be: a < aa < i < ii < ai < u < uu < au < e < o. Hence, the decision on how to encode the diphthongs is dependent upon how the vowels are handled.

(3) The six “resonant breathy” consonants are encoded as separate characters in L2/12-120, but are handled as conjuncts in L2/12-003. Based on Table 13, it appears that these “breathy” consonants are treated in collation as a unit, and not as a sequence. If this is so, encoding them as separate characters would be warranted.

It was noted that L2/12-003 also contains dependent vocalic characters L, LL, and RR, as well as ten additional signs and punctuation marks that appear in historic documents. While such historic characters do not appear in modern usage, documented use for these characters still makes them
eligible for encoding.

**Recommendation:** We recommend the UTC advise Anshuman Pandey take into account the information contained in Tables 11, 13 and 14, which includes analyses from the user community, and any comments from the UTC discussion, and revise his proposal accordingly.

**VII. BENGALI**  
**Document:** L2/12-121 Proposal to Encode the Sign ANJI for Bengali - Anshuman Pandey  
**Discussion:** We reviewed this document and found it to be sound.  
**Recommendation:** We recommend the UTC approve the character. We note that there will be ample opportunity for National Bodies to comment on it, once it is on a ballot.

**VIII. DEVANAGARI**  
**Document:** L2/12-123 Proposal to Encode the Sign SIDDHAM for Devanagari -Anshuman Pandey  
**Discussion:** We reviewed this document. Because the layout for Tibetan is quite different from Devanagari, the separate encoding of this character for Devanagari is warranted in our opinion.  
**Recommendation:** We recommend the character be renamed to “DEVANAGARI SIDDHAM SIGN”, in order to help differentiate it from script “Siddham”, and that the UTC approve it.

**IX. SHARADA**  
**Document:** L2/12-124 Proposal to Encode Signs for Writing Kashmiri in Sharada -Anshuman Pandey  
**Discussion:** We reviewed this proposal, and find it sound.  
**Recommendation:** We recommend the UTC accept these characters.

**X. WARANG CITI**  
**Document:** L2/12-118 Final proposal for encoding the Warang Citi script (WG2 N4259)  
**Discussion:** We reviewed this document, which now has received approval from an expert in the user community. The revised proposal also incorporates details on the handling of conjuncts, as was discussed in the February 2012 UTC.  
**Recommendation:** We recommend the UTC accept this script for encoding.

**XI. KAWI**  
**Document:** L2/12-125 Preliminary Proposal to Encode the Kawi Script – Anshuman Pandey  
**Discussion:** We reviewed this document. Unless there is more evidence in support of the character positions left open, we recommend the proposed block remove unnecessary spaces and packed down, thereby taking up only four columns, instead of five.  
**Recommendation:** We recommend the author do additional research to substantiate the request for
five columns, and keep the Roadmap Committee apprised. We also suggest UTC members review this document and send any comments to the proposal author.