TO: UTC
FROM: Deborah Anderson, Rick McGowan, Ken Whistler, and Roozbeh Pournader
SUBJECT: Recommendations to UTC on Script Proposals
DATE: 26 April 2013

In an effort to speed up processing of scripts in the UTC, we have reviewed the following script/character encoding proposals and provide recommendations. In general, we grouped them into one of the following categories:

- noncontroversial, and the UTC should proceed with approval
- unacceptable, and the UTC should go on record as not wanting to progress the proposal
- problematical, and the UTC needs to discuss on specifically identified points
- noted, and the UTC take no action.

Note: Only those documents available to us on 24 April were reviewed. Documents posted after this date are not included.

1. L2/13-054 Proposal to encode Duodecimal Digit Forms – Pentzlin – 2013-04-01

Discussion: We reviewed this proposal, which provided good text examples.

In our opinion, the RAISED CHI could be unified with A7B3 LATIN CAPITAL LETTER CHI, and the TURNED EZH added to the Latin Extended-D block. The other two characters, currently named "DIGIT FORM TWO" and "DIGIT FORM THREE," would be acceptable additions to the Number Forms block, which is the location proposed.

Note: The cross-reference suggested for TURNED EZH should be corrected to 01B7, not 0187 as in the proposal.

Recommendation: We recommend the UTC review the proposal and address the following questions:

a. Names
Are the names DIGIT FORM TWO and DIGIT FORM THREE appropriate? There are no other instances of DIGIT FORM as character names.

b. Identity of the characters:
Should RAISED CHI be unified with A7B3 LATIN CAPITAL LETTER CHI (see L2/12-270), adding an appropriate annotation?

c. Properties
The general category property proposed for all four characters is "No", but would "So" be more appropriate for TURNED TWO and TURNED THREE?

If RAISED CHI is unified with A7B3 LATIN CAPITAL LETTER CHI, it would have the "Lu" property. What property should TURNED EZH have?

2. [L2/13-055](#) Proposal to encode symbols for penalty cards – Pentzlin – 2013-04-01

Discussion: We reviewed this proposal, which provided no plain text justification for encoding the characters. Encoding such characters would, in our view, be problematical, since users would expect colors to appear. A more advisable approach would be to use images.

Recommendation: We recommend the UTC deem the proposed characters as unacceptable.

3. [L2/13-049](#) Declaration for declining the "Hungarian" block of the DAM – Tamás Rumi, et al –2013-03-11

Discussion: We reviewed this document.

Recommendation: We recommend the UTC note this feedback, but take no action.

4. [L2/13-070](#) Proposal to Encode the Rohingya Script – Pandey – 2013-04-23

Discussion: We reviewed this document. The encoding model seems appropriate, and the proposal seems to be very well worked out. In our opinion, Rohingya may be ready for encoding, after discussion in the UTC to see if there are any outstanding questions or issues. (For example, the confusability section may need careful review.)

Recommendation: We recommend the UTC discuss the proposal.

5. [L2/13-068](#) Proposal to Encode the Mongolian Square Script – Pandey –2013-04-23

Discussion: We reviewed this proposal. The encoding model chosen for Mongolian Square uses a virama as subjoiner for encoding stacks. As a result, the encoding for this long dead script is compact, and does not include subjoined characters.

The rationale in the proposal for this model seems reasonable. The proposal has been reviewed by Andrew West and Chris Fynn, who both find the model acceptable. (This is important because the script is used to transcribe Tibetan.)

The proposed location matches that already on the Roadmap for Mongolian Square.

Recommendation: We recommend the UTC discuss the proposal, in particular the encoding model.

6. [L2/13-069](#) Revised Proposal to Encode the Soyombo Script – Pandey –2013-04-23

Discussion: We reviewed this proposal, which adopts a different model than Mongolian Square, namely, it uses the Tibetan model for stacks. The advantage of the Tibetan model is that it makes the script easier to implement, since there are already working Tibetan implementations which could be tweaked to handle Soyombo.

Another approach suggested by A. Pandey (offline) would be to consider the atomic letter as the nucleus, instead of the atomic letter being the frame + nucleus. In this scenario, a user would enter the frame character and then the nucleus for each letter. Hence, instead of one character per letter, there would be two. The advantage of this approach is that writing of conjuncts would be easier, since users could place as many nuclei next to one another as needed, and thereby eliminate the need to encode a set of regular and subjoined characters for each consonant. However, it would require the two characters be typed for a single letter.

The current model, if approved as proposed, would require it be relocated on the Roadmap, since it would require two additional columns (for a total of 8). If the current proposal is approved, we suggest Soyombo be swapped with Ranjana (which currently covers 6 columns with 2 open columns beside it).

Recommendation: We recommend the UTC discuss the proposal, in particular the encoding model.

7. L2/13-067 Preliminary Code Chart for the Pau Cin Hau Syllabary – Pandey – 2013-04-23

Discussion: We reviewed this preliminary proposal. We note that the repertoire extends beyond the Roadmap allocation. However, the document states on page 1 that it currently includes characters made up of chunks of base and combining diacritic, which will be further analyzed, so the final size of the repertoire will likely be reduced. The author has also mentioned offline that the glyphs will be cleaned up in the next version.

Recommendation: We recommend the UTC review the proposal and send the author any comments.

8. L2/13-056 Proposal to Encode the Sign JAIN OM for Devanagari – Pandey –2013-04-23

Discussion: We reviewed this proposal for one character, which includes extensive evidence of its use.

Recommendation: We recommend the UTC approve DEVANAGARI JAIN OM.

9. L2/13-066 Gujarati Sign Triple Nukta – Vinodh Rajan –2013-04-23

Discussion: We reviewed this document, which clearly identifies the need to represent the Avestan letter ZHE in Gujarati.

There are two possible approaches: This proposal advocates representing Avestan ZHE by 0A9C GUJARATI LETTER JA and a new triple nukta combining character, which is parallel to the way other Avestan letters can presently be transliterated in Gujarati:

Avestan TTE: Gujarati TA + NUKTA
Avestan NGE: Gujarati ANUSVARA + GHA + NUKTA
Avestan ZE: Gujarati JA + NUKTA
Avestan KHVE: Gujarati+ NUKTA

The proposal does not mention whether other letters take the triple nukta. A possible issue is whether the combining mark's placement would be problematical.

A second approach is to encode a new precomposed character, as was done for 0979 DEVANAGARI LETTER ZHA.

Recommendation: We recommend the UTC discuss this proposal and decide whether to give greater precedence to the current approach of transliterating Avestan in Gujarati, which makes use of a combining character, or the Devanagari atomic-letter approach.

10. L2/13-065 The present proposal of Grantham in pipeline will clone classical language Sanskrit – N D Logasundaram – 2013-04-23

Discussion: We reviewed this document

Recommendation: We recommend the UTC note it, but take no action.

11. L2/13-061 Proposal to encode Grantha Anusvara Above – Sharma – 2013-04-09

Discussion: We reviewed this proposal, which makes a clear case to support the encoding of the GRANTHA ANUSVARA ABOVE character, with solid documentation.

Recommendation: We recommend the UTC approve 1137D GRANTHA SIGN COMBINING ANUSVARA ABOVE

12. L2/13-062 Proposal to encode Grantha OM – Sharma – 2013-04-09

Discussion: We reviewed this document, which provides clear evidence for this rare character. The documentation addresses the concerns expressed by the Government of India in L2/10-409, which stated there was no evidence for such a character and OM could simply be represented by O + ANUSVARA or O+M.

Recommendation: We recommend the UTC approve 11350 GRANTHA OM.

13. L2/10-341 Proposal to encode Grantha sub-base vocalic signs – Sharma – 2010-09-15
Ancillary document:
L2/11-351 E-mail regarding Grantha sub-base vocalic L/LL and Vedic Tone Asterisk Above – Constable – 2011-10-13

Discussion: We reviewed these documents and deem the issue problematical, requiring input from UTC members who have Indic implementations.

Recommendation: We recommend the UTC discuss these documents.

14. L2/13-047 Revised proposal to encode Tamil fractions and symbols – Sharma   – 2013-03-11

Discussion: We reviewed this proposal which has addressed concerns raised in the last script report (L2/13-028). (The proposal was forwarded by the author to INFITT, GOI, GOTN and ICTA Sri Lanka in early March.)

Recommendation: We recommend the UTC accept the characters, after discussion and review.

15. L2/13-051 Proposal to encode Malayalam minor fractions – Sharma – 2013-03-28

Discussion: We reviewed this document, which looks reasonable. The grouping of the three more common fractions in one set (but in the same order as occurs in the Oriya chart) seems well justified. The properties appear correct. If the UTC finds it acceptable, it would make sense to move this proposal in tandem with the Tamil fraction proposal.

Recommendation: We recommend the UTC accept the characters, after discussion and review.

16. L2/13-071 Proposal to encode the Uyghur script – Omarjan Osman –2013-04-23

Discussion: We reviewed this document, which shows clear improvement over the earlier version (L2/12-066). In our view, the fundamental issue remains, however, that is, whether the vertical characters are presentation forms of the horizontal characters, or are actually a separate script. We recommend Uyghur experts be consulted on the proposal, in particular on the vertical/horizontal topic as well as comments on the digits. The topic of which combining dots should be encoded should be addressed. It would seem advisable that work on this proposal be done at the same time as work on a Sogdian proposal.

We suggest the author work up a revised version of the proposal, concentrating on topics directly relevant to a script proposal for Old Uyghur – the history of the script and its relation to Aramaic, Sogdian, and Mongolian; a map of the area using the script; details on the script as mentioned above.

Recommendation: We recommend the UTC discuss the proposal and send comments to the author (perhaps via D. Anderson who has agreed to act as interlocutor).