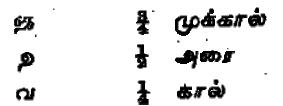
Request for encoding Tamil Major Fractions

Shriramana Sharma – jamadagni-at-gmail-dot-com 2009-Oct-25

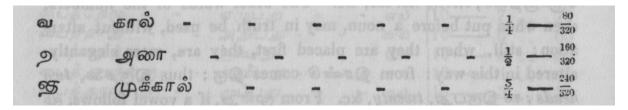
This is a request for encoding three characters in the Tamil block. These characters have historically been used in the Tamil script to denote the major fractions one quarter, one half and three quarters. The Grantha script also uses them for the same purpose, just as it uses the same digits 0-9 and numbers 10-1000 encoded in the Tamil block at 0BE6-0BF2.

The character for the fraction one quarter is glyphically very similar or identical to the Tamil (and Grantha) letter VA. The character for three quarters is glyphically very similar to the Grantha letter NYA. However, these characters should be encoded separately because when used to denote fractions they get different Unicode character properties.

The following sample from page 202 of ref 1, a Tamil grammar text which has been referenced by Unicode Technical Note #21 on Tamil Numbers, attests the usage of these fractions in the context of the Tamil script:



Ref 2 also shows these characters for Tamil:



The usage of these characters in Grantha is attested by the sample shown on the next page from page 273 of ref 4, which is a text on Vedic phonetics.

We note that while these characters are shown in the Grantha text placed above the mainline text, there is no need to encode separate above-combining versions of these characters because the desired effect can be achieved by ruby annotation. These fractions have been used in this Grantha text with their regular semantics, and denote the length in

mora-s of the vowel or consonant above which they are placed. Therefore they are unlike Sama Vedic svara markers which, though glyphically identical to digits placed above their base, do not have their regular semantics and instead mostly denote various svara-s.

Need for encoding in the Tamil block

I have previously submitted L2/09-317 requesting the allocation of a Tamil Supplementary block in the SMP to encode unified Grantha-Tamil fractions and other symbols. However, it would be appropriate to encode the three major fractions alone in the BMP in the Tamil block, just as Malayalam has its major fractions encoded in its own block at 0D73-0D75. The Tamil minor fractions may be separately encoded in the SMP, just as the international minor fractions are encoded at 2150-215E separate from the major fractions at 00BC-00BE.

Unicode character properties

```
OBFB; TAMIL FRACTION ONE QUARTER; No; 0; L;;;; 1/4; N;;;;;
OBFC; TAMIL FRACTION ONE HALF; No; 0; L;;;; 1/2; N;;;;
OBFD; TAMIL FRACTION THREE QUARTERS; No; 0; L;;;; 3/4; N;;;;
```

References

- 1. Iniya Tami<u>l</u> Ilakkaṇam, Yoki Śrī Cuttānanta Pāratiyār, Kavita Publications, T Nagar, Chennai; images from http://tamilelibrary.org/teli/numeral.html, 2009-Oct-21
- 2. http://kanimozhi.org.in/kanimozhi/?p=18, retrieved 2009-Oct-21
- 3. Vyāsa Śikṣā with Veda Taijasa commentary, Brahmananda Press, Tiruvadi, 1908.

Official Proposal Summary Form

A. Administrative

1 Title

Request for encoding Tamil Major Fractions

2. Requester's name

Shriramana Sharma

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution

4. Submission date

2009-Oct-25

- 5. Requester's reference (if applicable)
- 6. Choose one of the following:
- 6a. This is a complete proposal

Yes.

6b. More information will be provided later

No.

B. Technical - General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

No.

1b. Proposed name of script

1c. The proposal is for addition of character(s) to an existing block

Yes

1d. Name of the existing block

Tamil

2. Number of characters in proposal

3 (three)

3. Proposed category (A-Contemporary)

Category A.

4a. Is a repertoire including character names provided?

Yes

4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?

4c. Are the character shapes attached in a legible form suitable for review?

Yes

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Shriramana Sharma (for these characters alone).

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

$Shriramana\ Sharma,\ Font Forge.\ TTF\ file\ submitted\ along with\ proposal.$

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

Yes.

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script.

See detailed proposal.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

No.

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

Shriramana Sharma is part of the user community.

2b. If YES, with whom?

Vinodh Rajan and Sri Ramadoss M of Chennai.

2c. If YES, available relevant documents

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Tamil is spoken and written by a large user community in Tamil Nadu and elsewhere. The Grantha script is used for Sanskrit by Vedic scholars and others reading/writing Sanskrit in Tamil Nadu.

4a. The context of use for the proposed characters (type of use; common or rare)

Rare usage in contemporary Tamil books. Sometimes seen in Sanskrit works in Grantha.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

Tamil Nadu, India.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Yes.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

The rationale is that related characters are encoded in the BMP and there is space in the block.

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Yes. The reason is that they denote the three major fractions and are preferably kept together.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

The character for quarter is similar to Grantha/Tamil Letter VA. That for three quarters is similar to Grantha Letter NYA.

10b. If YES, is a rationale for its inclusion provided?

Yes.

10c. If YES, reference

Their Unicode character properties would differ.

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

No.

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

Nο.

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.