Proposal to encode 11350 GRANTHA OM

Shriramana Sharma, jamadagni-at-gmail-dot-com, India

2013-Apr-10

L2/13-062

§1. Background

The GoI committee on Grantha has previously noted in L2/10-409 p 4 that there is no evidence for a Grantha OM and that in Grantha, OM is usually only written as a sequence of O + ANUSVARA or O + M. Therefore no Grantha OM was recommended to be encoded.

In other previous documents L2/09-316 p 18 §IV, L2/10-267 p 1 and L2/10-286 also it was noted that the glyph which was previously proposed by other parties for the Grantha OM (and hence was present in L2/10-265 but removed in 265R) has not been properly attested to be the Grantha OM and hence it was insisted that the codepoint for Grantha OM be left reserved pending proper attestation in Grantha.

§2. New evidence for Grantha OM

New evidence has been uncovered for a distinct written form of OM in Grantha. See the following sample from a manuscript-style printing of a Sanskrit grammar work in Grantha script, obtained from a private collection near Chennai:
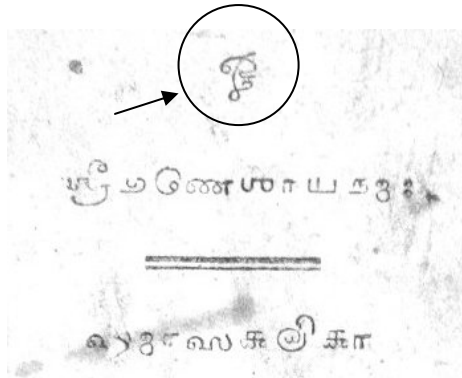
The information superscribed on the cover gives the date as 1-1-1910 and former owner as one Ramakrishna Sastriar of Villianalloor, a village near Kumbakonam, Tamil Nadu. This book is thus over a century old. A digital copy of this book has also been uploaded to http://archive.org/details/SamasaKalika. It is to be noted that the Grantha script seen in this book is of a very old style and exhibits archaic vowel sign ligatures, archaic consonant ligatures etc, confirming its authenticity.

This evidence has been presented to Grantha manuscript scholars in Tamil Nadu and elsewhere. Manuscripts scholars consulted by us have independently confirmed the identity of this character as the Grantha OM and have noted that it is very rare but definitely seen in old manuscripts in the Grantha script. (Contact information for above manuscript scholars is available upon request.) It is hence strongly indicated that other old manuscripts/prints using the same or similar glyph for OM may also be found.

It is noted that this glyph is a ligature of Grantha O + Grantha vowelless-MA:

$$\text{ɛɔ} + \text{ɞ} = \text{ɞ}$$

Thus it is appropriate to recognize this as the Grantha OM and hence encode it at the reserved codepoint 11350 with the glyph as shown in the attestation sample.

§3. Justification for encoding

The 2011 Aug UTC meeting noted that an important criteria in whether ligated forms of OM should be encoded separate from non-ligated forms or not is whether the native users of the script require the ligated form to be recognized as equal to the unligated form in text processes such as searching, or whether it should be treated as distinct.

In the case of Tamil, it was found that the ligated form should be treated distinctly and hence it was encoded. In the case of Oriya, the native users decided otherwise, as it is a simple ligature of O + CANDRABINDU within a grapheme cluster. (See L2/11-343.) However, the Grantha situation is parallel to Tamil, with the ligation taking place across two grapheme clusters. Native users therefore feel it should be identified distinctly in searching. Hence separate encoding is requested for Grantha.

§4. Character to be encoded

11350   GRANTHA OM

§5. Unicode Character Properties

```
11350;GRANTHA OM;Lo;0;L;;;;;N;;;;;
```

2

§6. Official Proposal Summary Form

(Based on N3902-F)

A. Administrative

1. Title
Proposal to encode 11350 GRANTHA OM
2. Requester's name
Shriramana Sharma
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution
4. Submission date
2013-Apr-10
5. Requester's reference (if applicable)
6. Choose one of the following: This is a complete proposal (or) More information will be provided later
This is a complete proposal.

B. Technical – General

1. Choose one of the following:
1a. This proposal is for a new script (set of characters), Proposed name of script
No.
1b. The proposal is for addition of character(s) to an existing block, Name of the existing block
Yes. Grantha.
2. Number of characters in proposal
1 (one)
3. Proposed category
Category B1, specialized small (for this character, though Grantha itself is "A, contemporary")
4. Is a repertoire including character names provided?
Yes.
4a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?
Yes.
4b. Are the character shapes attached in a legible form suitable for review?
Yes.
5. Fonts related:
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
Shriramana Sharma.
b. Identify the party granting a license for use of the font by the editors (include address, e-mail etc.)
Shriramana Sharma.
6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
No.
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)?
Not applicable.
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.)?
Yes.
2b. If YES, with whom?

Dr Sheshadri Ghanapathi, Retd from the Government Oriental Manuscripts Library, Chennai. Dr Shankaranarayanan, Sanskrit Dept, Chief of Manuscripts Section, SCSVMV University, Kanchipuram. Dr R Krishnamurti Shastri, Retd Principal, Madras Sanskrit College, Chennai. Dr Kirtikant Sharma, Manuscript Unit, IGNCA, New Delhi. Dr T Ganesan, Manuscripts scholar, IFP, Pondicherry.

2c. If YES, available relevant documents

None specifically. The matter was discussed in person and via email/phone.

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

Manuscript scholars desiring to store contents of Grantha manuscripts as e-text

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

Rare.

4b. Reference

See detailed proposal.

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

It would be mostly used only by manuscript scholars.

5b. If YES, where?

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

No.

6b. If YES, is a rationale provided?

6c. If YES, reference

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

Only one character is proposed. It should be placed in the codepoint reserved for it.

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. Though it is visually a ligature of O + vowelless-MA, native users would like it to be treated separately from the sequence of O + vowelless-MA. Thus a separate character is needed.

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?

No.

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

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences?

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?

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.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

13c. If YES, reference:

–o–o–o–