Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: **Working Group Document**
Title: **Proposal to add two characters for Malayalam to the BMP of the UCS**
Source: **Michael Everson**
Status: **Individual Contribution**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Replaces: **N3295 (L2/07-283)**
Date: **2007-08-14**

This document requests two additional characters to be added to the UCS and contains the proposal summary form. The characters are as follows:

0D29   MALAYALAM LETTER NNNA
0D3A   MALAYALAM LETTER TTTA

These characters represent alveolar consonants, transliterated $\underline{n}$ and $\underline{t}$. They are used in scholarly orthography which distinguishes the alveolar sounds. The first has an analogue in modern Tamil, U+0B29 TAMIL LETTER NNNA. The second is another innovation which was subsequently lost because it merged with U+0D31 MALAYALAM LETTER RRA; RRA is pronounced as aleolar $\underline{r}$ or alveolar $\underline{t}$ depending upon context in modern Malayalam orthography.

It is also proposed that the font currently used in the Malayalam code chart be replaced by the font used in this document. This font has better proportions and, for instance, does not suffer from the many weight problems of the font currently in use; compare the glyph in U+0D6B for instance.

Unicode Character Properties

```
0D29;MALAYALAM LETTER NNNA;Lo;0;L;;;;;N;;;;;
0D3A;MALAYALAM LETTER TTTA;Lo;0;L;;;;;N;;;;;
```

Bibliography

Grünendahl, Reinhold. 2001. *South Indian scripts in Sanskrit manuscripts and prints: Grantha Tamil, Malayalam, Telugu, Kannada, Nandinagari*. Wiesbaden: Harrassowitz. ISBN 3-447-04504-3
Pedersen, Thomas T. 2005-09-30. *Transliteration of Malayalam*. Retrieved 2007-08-11 from
   http://transliteration.eki.ee/pdf/Malayalam.pdf
Rajaraja Varma, A. R. 1917. *Keralapanineeyam*. Reprinted 2003. Kottayam: DC Books.

Figures

Figure 1. Sample from Grünendahl 2001 showing NNNA.

Figure 2. Sample from Pedersen 2005 showing NNNA and TTTA.

Figure 3. Sample from Rajaraja Varma 1917 showing NNNA and TTTA.

Row 0D: MALAYALAM

	0D0	0D1	0D2	0D3	0D4	0D5	0D6	0D7
0		ഐ	ഠ	ര	ീ		ൌ	ൠ
1			ഡ	റ	ു		ആ	ൡ
2	ം	ഒ	ഢ	ല	ൂ		�	ൢ
3	ഃ	ഓ	ണ	ള	ൃ		�	ൟ
4		ഔ	ത	ഴ	ൄ			ൣ
5	അ	ക	ഥ	വ				൥
6	ആ	ഖ	ദ	ശ	െ		൦	
7	ഇ	ഗ	ധ	ഷ	േ	ൗ	൧	
8	ഈ	ഘ	ന	സ	ൈ		൨	
9	ഉ	ങ	ണ	ഹ			൩	൯
A	ഊ	ച	പ	ട	ൊ		൪	ൺ
B	ഋ	ഛ	ഫ		ോ		൫	ൻ
C	ഌ	ജ	ബ		ൌ		൬	ർ
D		ഝ	ഭ	ഺ	്		൭	ൽ
E	എ	ഞ	മ	ാ			൮	ൾ
F	ഏ	ട	യ	ി			ൻ	ൿ

Row 0D: MALAYALAM

hex	Name
00	(This position shall not be used)
01	(This position shall not be used)
02	MALAYALAM SIGN ANUSVARA
03	MALAYALAM SIGN VISARGA
04	(This position shall not be used)
05	MALAYALAM LETTER A
06	MALAYALAM LETTER AA
07	MALAYALAM LETTER I
08	MALAYALAM LETTER II
09	MALAYALAM LETTER U
0A	MALAYALAM LETTER UU
0B	MALAYALAM LETTER VOCALIC R
0C	MALAYALAM LETTER VOCALIC L
0D	(This position shall not be used)
0E	MALAYALAM LETTER E
0F	MALAYALAM LETTER EE
10	MALAYALAM LETTER AI
11	(This position shall not be used)
12	MALAYALAM LETTER O
13	MALAYALAM LETTER OO
14	MALAYALAM LETTER AU
15	MALAYALAM LETTER KA
16	MALAYALAM LETTER KHA
17	MALAYALAM LETTER GA
18	MALAYALAM LETTER GHA
19	MALAYALAM LETTER NGA
1A	MALAYALAM LETTER CA
1B	MALAYALAM LETTER CHA
1C	MALAYALAM LETTER JA
1D	MALAYALAM LETTER JHA
1E	MALAYALAM LETTER NYA
1F	MALAYALAM LETTER TTA
20	MALAYALAM LETTER TTHA
21	MALAYALAM LETTER DDA
22	MALAYALAM LETTER DDHA
23	MALAYALAM LETTER NNA
24	MALAYALAM LETTER TA
25	MALAYALAM LETTER THA
26	MALAYALAM LETTER DA
27	MALAYALAM LETTER DHA
28	MALAYALAM LETTER NA
29	MALAYALAM LETTER NNNA
2A	MALAYALAM LETTER PA
2B	MALAYALAM LETTER PHA
2C	MALAYALAM LETTER BA
2D	MALAYALAM LETTER BHA
2E	MALAYALAM LETTER MA
2F	MALAYALAM LETTER YA
30	MALAYALAM LETTER RA
31	MALAYALAM LETTER RRA
32	MALAYALAM LETTER LA
33	MALAYALAM LETTER LLA
34	MALAYALAM LETTER LLLA
35	MALAYALAM LETTER VA
36	MALAYALAM LETTER SHA
37	MALAYALAM LETTER SSA
38	MALAYALAM LETTER SA
39	MALAYALAM LETTER HA
3A	MALAYALAM LETTER TTTA
3B	(This position shall not be used)
3C	(This position shall not be used)
3D	MALAYALAM SIGN AVAGRAHA (praslesham)
3E	MALAYALAM VOWEL SIGN AA
3F	MALAYALAM VOWEL SIGN I
40	MALAYALAM VOWEL SIGN II
41	MALAYALAM VOWEL SIGN U
42	MALAYALAM VOWEL SIGN UU
43	MALAYALAM VOWEL SIGN VOCALIC R
44	MALAYALAM VOWEL SIGN VOCALIC RR
45	(This position shall not be used)
46	MALAYALAM VOWEL SIGN E
47	MALAYALAM VOWEL SIGN EE
48	MALAYALAM VOWEL SIGN AI
49	(This position shall not be used)
4A	MALAYALAM VOWEL SIGN O
4B	MALAYALAM VOWEL SIGN OO
4C	MALAYALAM VOWEL SIGN AU
4D	MALAYALAM SIGN VIRAMA
4E	(This position shall not be used)
4F	(This position shall not be used)
50	(This position shall not be used)
51	(This position shall not be used)
52	(This position shall not be used)
53	(This position shall not be used)
54	(This position shall not be used)
55	(This position shall not be used)
56	(This position shall not be used)
57	MALAYALAM AU LENGTH MARK
58	(This position shall not be used)
59	(This position shall not be used)
5A	(This position shall not be used)
5B	(This position shall not be used)
5C	(This position shall not be used)
5D	(This position shall not be used)
5E	(This position shall not be used)
5F	(This position shall not be used)
60	MALAYALAM LETTER VOCALIC RR
61	MALAYALAM LETTER VOCALIC LL
62	MALAYALAM VOWEL SIGN VOCALIC L
63	MALAYALAM VOWEL SIGN VOCALIC LL
64	(This position shall not be used)
65	(This position shall not be used)
66	MALAYALAM DIGIT ZERO
67	MALAYALAM DIGIT ONE
68	MALAYALAM DIGIT TWO
69	MALAYALAM DIGIT THREE
6A	MALAYALAM DIGIT FOUR
6B	MALAYALAM DIGIT FIVE
6C	MALAYALAM DIGIT SIX
6D	MALAYALAM DIGIT SEVEN
6E	MALAYALAM DIGIT EIGHT
6F	MALAYALAM DIGIT NINE
70	MALAYALAM NUMBER TEN
71	MALAYALAM NUMBER ONE HUNDRED
72	MALAYALAM NUMBER ONE THOUSAND
73	MALAYALAM FRACTION ONE QUARTER
74	MALAYALAM FRACTION ONE HALF
75	MALAYALAM FRACTION THREE QUARTERS
76	(This position shall not be used)
77	(This position shall not be used)
78	(This position shall not be used)
79	MALAYALAM DATE MARK
7A	MALAYALAM LETTER CHILLU NN
7B	MALAYALAM LETTER CHILLU N
7C	MALAYALAM LETTER CHILLU RR
7D	MALAYALAM LETTER CHILLU L
7E	MALAYALAM LETTER CHILLU LL
7F	MALAYALAM LETTER CHILLU K

A. Administrative

1. Title
Proposal for to add two characters for Malayalam to the BMP of the UCS.
2. Requester's name
Michael Everson
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution.
4. Submission date
2007-08-14
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
Malayalam.
2. Number of characters in proposal
2.
3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)
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?
Yes.
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?
Michael Everson.
5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Michael Everson, Fontographer.
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. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.
See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
Yes, in N3295.
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?
C. J. Cibu. Naga Ganesan

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?

Malayalam users.

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

Rare.

4b. Reference

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

Yes.

5b. If YES, where?

Dictionaries, grammars, books on linguistics.

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?

No.

6c. If YES, reference

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

No.

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?

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 (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?

No.

11e. If YES, reference

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?