

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹**

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

| | |
|--|------------------------------------|
| 1. Title: | <i>Malayalam Ordinal Indicator</i> |
| 2. Requester's name: | <i>Håkan Davidsson</i> |
| 3. Requester type (Member body/Liaison/Individual contribution): | <i>Individual contribution</i> |
| 4. Submission date: | <i>December 16, 2014</i> |
| 5. Requester's reference (if applicable): | <i>N/A</i> |
| 6. Choose one of the following: | |
| This is a complete proposal: | <i>Yes</i> |
| (or) More information will be provided later: | <i>-</i> |

B. Technical – General

| | |
|---|--|
| 1. Choose one of the following: | |
| a. This proposal is for a new script (set of characters): | <i>N/A</i> |
| Proposed name of script: | <i>N/A</i> |
| b. The proposal is for addition of character(s) to an existing block: | <i>Yes</i> |
| Name of the existing block: | <i>Malayalam 0D00-0D7F</i> |
| 2. Number of characters in proposal: | <i>One</i> |
| 3. Proposed category (select one from below - see section 2.2 of P&P document): | |
| A-Contemporary <input type="checkbox"/> B.1-Specialized (small collection) <input type="checkbox"/> B.2-Specialized (large collection) <input type="checkbox"/> | |
| C-Major extinct <input type="checkbox"/> D-Attested extinct <input type="checkbox"/> E-Minor extinct <input type="checkbox"/> | |
| F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/> G-Obscure or questionable usage symbols <input type="checkbox"/> | |
| 4. Is a repertoire including character names provided? | <i>Yes</i> |
| a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? | <i>Yes</i> |
| b. Are the character shapes attached in a legible form suitable for review? | <i>Yes</i> |
| 5. Fonts related: | |
| a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard? | <i>We have enclosed a FontLab source font. Likely Microsoft, Apple etc. will include this glyph in their fonts</i> |
| b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): | <i>N/A</i> |
| 6. References: | |
| a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? | <i>Yes</i> |
| b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? | <i>Yes</i> |
| 7. Special encoding issues: | |
| 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)? | <i>Yes,</i> |
| | <i>The description and behavior can be the same as the Date Mark. It can be sorted before the Date Mark</i> |

8. Additional Information:

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/reports/tr44/>) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

¹ Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

| | |
|--|---|
| 1. Has this proposal for addition of character(s) been submitted before? If YES explain | No |
| 2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? If YES, with whom? If YES, available relevant documents: | Yes <i>State Institute of Languages in Trivandrum</i> No |
| 3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference: | Yes Link to iw.org web site |
| 4. The context of use for the proposed characters (type of use; common or rare) Reference: | Common <i>See link above</i> |
| 5. Are the proposed characters in current use by the user community? If YES, where? Reference: | Yes <i>All major Newspapers for example but used incorrectly 0D3E+0D02</i> |
| 6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference: | Yes Yes <i>See Addendum below</i> |
| 7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? | N/A |
| 8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference: | Yes Yes <i>See Addendum below</i> |
| 9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference: | No N/A N/A |
| 10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? If YES, is a rationale for its inclusion provided? If YES, reference: | No N/A N/A |
| 11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference: | No N/A N/A N/A |
| 12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary) | No N/A |
| 13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference: | No N/A N/A |

ADDENDUM

Proposal to include Malayalam Ordinal Indicator symbol

The Malayalam ordinal number symbol is not in Unicode. Unicode does have two dependent symbols: ഓ (0D3E - MALAYALAM VOWEL SIGN AA) and ഓ (0D02 - MALAYALAM SIGN ANUSVARA). Although these two characters can be combined so that they have the same appearance as the ordinal number symbol, they are both combining marks and therefore need a base code. There is no base code that is a natural fit for this purpose. Even if a zero-width joiner is used as a base code, they do not combine properly and do not display properly in all web browsers or word processing programs. For example Internet Explorer 10 displays these UTF-16 code points: 0031(1) 002D(-) 200D(ZWJ) 0D3E(ഓ) 200D(ZWJ) 0D02(ഓ) 1-ഓ as (which has the correct appearance) whereas Chrome version 37 renders the same as 1-ഓഓഓ . Safari also does not display this properly. Malayalam uses the ordinal number symbol ഓ in its written form much the same as English uses 1st 2nd 3rd 4th and so on.

Clearly, it would be useful to have a single Unicode symbol for the Malayalam Ordinal Number Symbol. This should be a spacing character and not a combining mark. This would allow for consistent rendering and text searching. Our website www.jw.org/ml is using ഓ in many places. However, it does not display properly in most web browsers. Many other organizations would also benefit if this proposal is supported and approved. We will also submit this proposal to the State Institute of Languages in Trivandrum.

By reading the following link it appears that our proposed "Malayalam Ordinal Indicator" has been in the pipeline before: <http://www.unicode.org/mail-arch/unicode-ml/y2006-m04/0180.html>. But we found no reason that indicates that this character was disapproved.

May we propose that this new character is added under the heading "Ordinal Indicator" using Unicode 0D77. This "Ordinal Indicator" can be rendered the same as the ഓ (0D3E - MALAYALAM VOWEL SIGN AA) and ഓ (0D02 - MALAYALAM SIGN ANUSVARA) which makes it render this way ഓ .

Here is a link to where we use this character:

<http://www.jw.org/ml/%E0%B4%AF%E0%B4%B9%E0%B5%8B%E0%B4%B5%E0%B4%AF%E0%B5%81%E0%B4%9F%E0%B5%86-%E0%B4%B8%E0%B4%BE%E0%B4%95%E0%B5%8D%E0%B4%B7%E0%B4%BF%E0%B4%95%E0%B5%BE/%E0%B4%AA%E0%B5%8D%E0%B4%B0%E0%B4%B5%E0%B5%BC%E0%B4%A4%E0%B5%8D%E0%B4%A4%E0%B4%A8%E0%B4%99%E0%B5%8D%E0%B4%99%E0%B5%BE/%E0%B4%AA%E0%B5%8D%E0%B4%B0%E0%B4%A4%E0%B5%8D%E0%B4%AF%E0%B5%87%E0%B4%95-%E0%B4%AA%E0%B4%B0%E0%B4%BF%E0%B4%AA%E0%B4%BE%E0%B4%9F%E0%B4%BF%E0%B4%95%E0%B5%BE/%E0%B4%B5%E0%B5%80%E0%B4%A1%E0%B4%BF%E0%B4%AF%E0%B5%8B-%E0%B4%97%E0%B4%BF%E0%B4%B2%E0%B5%86%E0%B4%AF%E0%B4%BE%E0%B4%A6%E0%B5%8D%E2%80%8C-%E0%B4%95%E0%B5%8D%E0%B4%B2%E0%B4%BE%E0%B4%B8%E0%B5%8D%E2%80%8C-135/>

Please note that these characters do not display properly in most browsers even if it does in IE 10. That also applies to Word and other similar programs which do not display this correctly.

We also enclose a font sample in the Font Lab source file of how this combination can be rendered.

Font - Malayalam-Ordinal-Indicator [D:\...\Unicode Submitting of Ordinal Indicator\Malayalam Ordinal Indicator.vfb]

Glyph - _0001 from Malayalam-Ordinal-Indicator

100 %

| | | | | | | | | | | | | | | | | | | |
|-------|-----|-----|-----|-------|-----------|-----|------|-----|----------|--------|-------|----------|-----|-----|-----|-----|-----|-----|
| _0001 | --- | --- | --- | breve | dotaccent | --- | ring | --- | hungarum | ogonek | caron | dotlessi | --- | --- | --- | --- | --- | --- |
| oo | | | | ˘ | ˙ | | ˚ | | ˛ | ˇ | ˆ | ̂ | | | | | | |