

L2/09-245R

**ISO/IEC JTC 1/SC 2/WG 2  
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS  
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646<sup>1</sup>**

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

Please read Principles and Procedures Document (P & P) from <http://www.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://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

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

**A. Administrative**

1. Title:	<i>Proposal to encode Malayalam Dot-Reph character</i>
2. Requester's name:	<i>US National Body</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Member body</i>
4. Submission date:	<i>2009 May 1</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<i>Yes</i>
(or) More information will be provided later:	<i>No</i>

**B. Technical – General**

1. Choose one of the following:		
a. This proposal is for a new script (set of characters):	<i>No</i>	
Proposed name of script:		
b. The proposal is for addition of character(s) to an existing block:	<i>Yes</i>	
Name of the existing block:	<i>Malayalam</i>	
2. Number of characters in proposal:	<i>1</i>	
3. Proposed category (select one from below - see section 2.2 of P&P document):		
A-Contemporary <input checked="" 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. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?		
If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:		
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>Suggested character properties and sort orders are included.</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 <http://www.unicode.org/Public/UNIDATA/UCD.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

<sup>1</sup> Form number: N3152-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)

### 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 Government of India L2/08-411, L2/08-306
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	See section D
4. The context of use for the proposed characters (type of use; common or rare) Reference:	rare See section A
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	Yes See section D
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  If possible, should be kept in the BMP since the rest of the script is in the BMP
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	Yes
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 See section B
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
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:	No
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:	Yes Yes See section C Yes See section F
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
13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? If YES, reference:	No

## A. Introduction

Dot-Reph is used to represent Malayalam letter RA\_dead, as a dot over the consonant following it. That is, in the sequence <Malayalam Akshara X, Malayalam consonant R, Akshara Y>, the consonant 'R' is represented as a dot over Akshara Y. As of Unicode 5.1, there is no representation for Dot-Reph.

## B. Rationale for encoding

The sequence <RA, VIRAMA, ZWJ> would have been the natural choice for Dot-Reph as in Indic scripts like Gurmukhi. However, that sequence has considerable existing usage, predating Unicode 5.1, to represent MALAYALAM CHILLU RR.

We could have opted for <RA, VIRAMA, Consonant> to be rendered with Dot-Reph while <RA, ZWJ, VIRAMA, Consonant> displaying RA with C2-conjoining form of the other consonant. For example, cluster ര്യ (/rya/) represented as <RA, ZWJ, VIRAMA, YA>.

However, sequences like <RA, VIRAMA, YA> is currently being used for the cluster ര്യ (/rya/) in very common words like സൂര്യൻ (/sooryan/ meaning 'sun') and ഭാര്യ (/bhaarya/ meaning 'wife'). Changing the rendering for those common sequences would be disastrous.

Another potential sequence is <RRA, VIRAMA> with or without a joiner. However, in Malayalam, the underlying base character for Dot-Reph is almost always <RA>. That would prevent us from using any sequence with <RRA> for Dot-Reph.

A sequence involving <CHILLU RR> and joiners would be particularly attractive choice since a native user understands Dot-Reph that way. However, that does not always hold good. For example, in very common words like സൂര്യൻ (/sooryan/ meaning 'sun') and ഭാര്യ (/bhaarya/ meaning 'wife'), their respective forms without Dot-Reph does not involve Chillu-RR. They are respectively സൂര്യൻ and ഭാര്യ.

Given the above observations, atomically encoding Dot-Reph may be the only plausible choice. Conceptually, Dot-Reph is analogous to C1-conjoining cluster <RA, Virama> (but not interchangeable with). However, the Malayalam cluster <RA, Virama, C2> can assume C2-conjoining form. For example: ര്യ (/rya/) cluster in the word ഭാര്യ (/bhaarya/) meaning wife. When C1-conjoining form needs to be specified - that is, the Dot-Reph form - the proposed Dot-Reph character needs to be used as <DOT-REPH, C2>.

Common Malayalam input keyboards would need to add one more key to support this character. It could be positioned in the third or fourth plane since its contemporary usage is minimal.

## C. Usage

Dot-Reph would be a spacing character and would be analogous to C1-conjoining <RA, VIRAMA> (but not interchangeable with). It would get displayed as a dot over the subsequent Malayalam consonant or cluster. Example:



<DOT-REPH, VA>

**D. Demographics and usage**

Dot-Reph was in common print usage until 1970. After that its usage has receded from print. Currently it is nowhere to be seen in contemporary publishing. The words that were using Dot-Reph are now transcribed using Chillu-RR or respective C2-conjoining forms. Those who have completed their elementary education in Malayalam prior to 1970, predominantly use this form for writing. Total number of Malayalam users is more than 35 million (1997) (Reference: [Ethnologue](#)).

**E. Character Name, position, collation and properties**



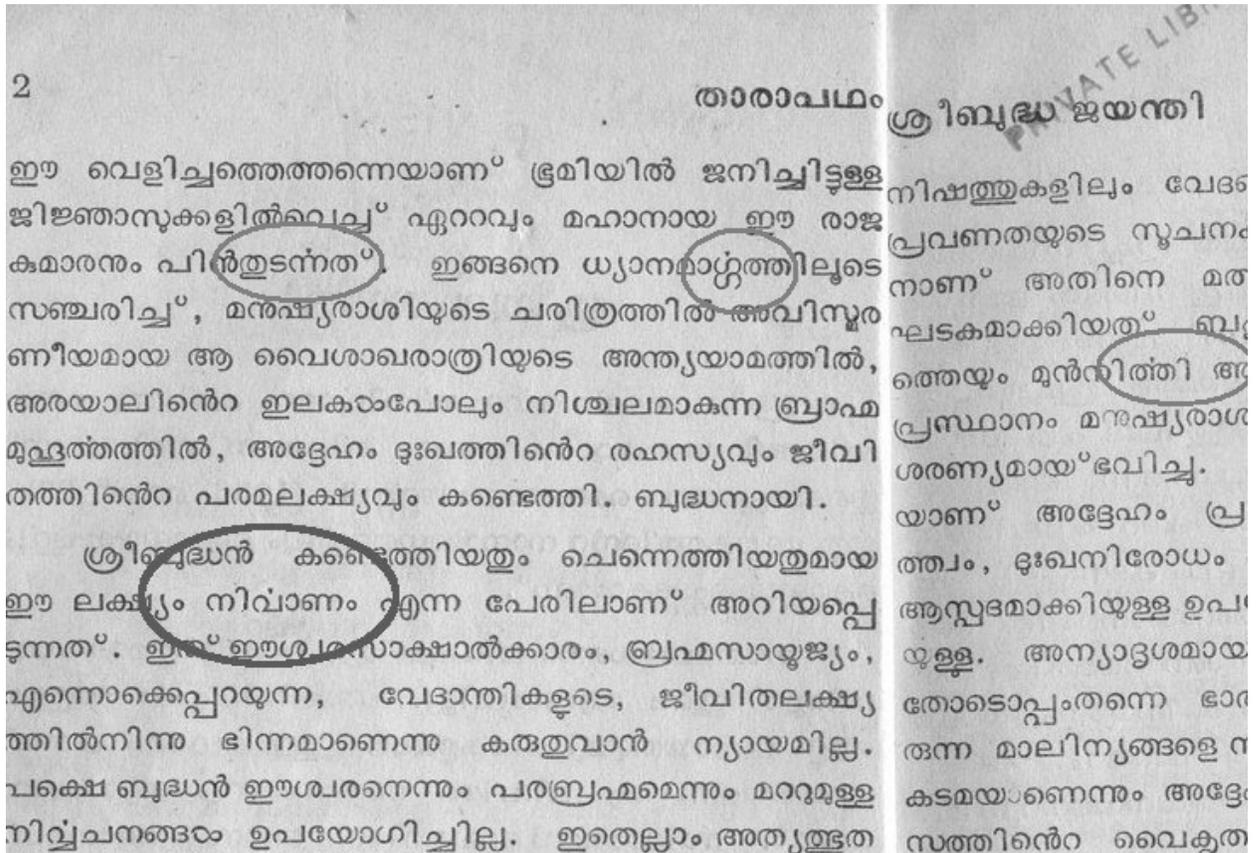
OD4E MALAYALAM LETTER DOT REPH

For collation, DOT REPH could be used placed with primary weight equal to Chillu-RR and secondary weight less than that.

Properties:

OD4E;MALAYALAM LETTER DOT REPH;Lo;0;L;:::;N;:::;

**F. Scans**



Malayalam language textbook for the grade 10; year 1966; published by Government of Kerala.

പാർവ്വണം pārvanam S. Connected with പ  
ർവ്വം, as പാർവ്വണേ the full-moon.

പാർവ്വതി pārvati S. (പാർവ്വതം) = മലമാതൃ; N.pr. f.

പാർവ്വത്യം pārvatyam (Tdbh. of പ്രാവൃത്യം? as  
if fr. പ്രവൃത്തി q. v.; but T. Te. V1. have pā-  
rapatyam, C. pā A Malayalam English Dictionary  
By Hermman Gundert, 1872 -  
nagement, as if fr. പാക്ക). 1. Local govern-  
ment, administration മോർ വില്ലുന്ന തായേ ഉര  
രിലേ പാരവത്യം എന്തിനു prov., also പാറോ  
ത്യം; നാട്ടിൽ കാരോരുത്തർ പാറോവത്തി ചെ  
യ്തു, പാരവത്യം ചെയ്യുന്നവൻ TR. 2. Trav.  
a subordinate revenue situation.

പാർവ്വത്യക്കാരൻ a subordinate Collector & Ma-  
gistrate (Trav. = അധികാരി). സൂലാൻറെൻ  
പാർവ്വത്യക്കാരന്മാർ, പാർവ്വതിക്കാരൻ TR. (In  
Mal. a district about as large as 4 amśams  
(അംശം 2) together, was formerly under 1  
പാർവ്വത്യക്കാരൻ: f. i. പാർവ്വതിക്കാർ പണ  
ത്തിനു വരാറായി No. said only by very  
old people).

പാർശ്വാപ് P. Pādshāh, The Mogul emperor