

Malayalam/Comparison of Chillu encoding proposals

From Unicode discussion

< Malayalam

Cibu C. Johny
cibu@google.com
2007-01-30

1 Introduction

Please read following background documents.

- n Malayalam/Representation of Vowellessness
- n L2/06-207 on the proposed solution for issues in Malayalam Chillu encoding

2 Definitions

- n 'ID' represents the systems where format control characters are stripped
- n 'IDN' is a ID system with additional restrictions.
- n 'Text' means sequence of Unicode characters without any filtering or restriction.

3 Comparison table

Option#	Option 1	Option 2	Option 3	Option 4
Proposals:	<Consonant, VIRAMA, ZWJ>	<Consonant, VIRAMA, special joiner>. This special Joiner ignored in some text processing like IDN and preserved in rest.	<Consonant, Chillu Sign>	Atomic encoding of Chillus
Current Status	Current standard	Discussed in the Indic list	Rejected	Proposed new standard
1. Can represent Chillu of ൠ (RRA)	No	No	No	Yes
2. Has a non-transparent representation of Chillu	No for IDN (words with Chillu cannot be registered as IDN)	No for IDN (words with Chillu cannot be registered as IDN)	Yes	Yes


3. Words like ന്ന and ന്ന that are visually different, but meaning the same, can be encoded the same	Yes; barring format control characters (ന്ന cannot be used in IDN)	No in text; Yes in IDN (ന്ന cannot be used in IDN)	No (needs higher level logic to establish the equivalence)	No (needs higher level logic to establish the equivalence)
4. If a new chillu letter is discovered, no need to encode it specifically	Yes in text; no chillu representation available for IDN or ID	Yes in text and ID; no chillu representation available in IDN	Yes	No
5. ണ് (/nta/) does not need a non-intuitive character sequence	False: <NA, VIRAMA, ZWJ, VIRAMA, RRA> with 2 consecutive VIRAMAs	False: <NA, VIRAMA, special joiner, VIRAMA, RRA>	Partly True: <NA, chillu-sign, VIRAMA, RRA>. Two consecutive VIRAMAs not required; However, equates ണ് to ണ്.	Partly True: <Chillu-NA, VIRAMA, RRA>. Two consecutive VIRAMAs not required; However, equates ണ് to ണ്.
6. Chillu-RA/RRA (റ) is intuitive to native user	No. Native user assumes that റ is based on RRA	No. Native user assumes that റ is based on RRA	No. Native user assumes that റ is based on RRA	Yes. Base character is not relevant at the encoding level

Retrieved from "http://unicode.wikia.com/wiki/Malayalam/Comparison_of_Chillu_encoding_proposals"

n
n This page was last modified 2007-01-30 21:22:06.

n [About Wikia](#)
n [Terms of use](#)
n [Advertise](#)

n 

n 

Wikia is a service mark of Wikia, Inc. All rights reserved.