

Unicode Standard for Grantha Script

Department of Information Technology,
Ministry of Communication & Information Technology, Government Of India

2010-Oct-18

Summary of the Meeting

New Delhi, September 6th, 2010

The Department of Information Technology, Government of India is a voting member of the Unicode Consortium for adequate representation of the Indian scripts in the Unicode Standard. The major scripts of all the 22 constitutionally recognized Indian languages have been represented in the Unicode Standard. Initiatives were taken to encode Indian heritage scripts i.e. Vedic Sanskrit and Grantha in the Unicode Standard so that our ancient knowledge could be represented on electronic media, computers etc.

Proposals were submitted to Unicode Technical Committee (UTC) for encoding of the Grantha script in the Unicode Standard. After the initial round of discussion UTC has in principle recommended encoding 90 characters as per the registry document L2/10-265R. These characters must be considered for finalized encoding in the standard after discussion and consensus with all the stakeholders.

A meeting of Grantha experts and other stakeholders was organized on September 6th, 2010 at the Department of Information Technology, 6 CGO Complex, Lodhi Road, New Delhi to discuss the various issues with respect to Grantha encoding and finalize the recommendation for onward submission to the Unicode Technical Committee. The following experts participated in the meeting:

1. Shri N Ravi Shanker, Joint Secretary, Deptt. Of IT, New Delhi;
2. Ms Swaran Lata, Deptt. Of IT, New Delhi;
3. Prof Brij Bihari Chaubey, Hoshiyarpur;
4. Shri Vrindavan Goswami, Vrindavan;
5. Dr R Krishnamurti Shastri, Heritage India Educational Trust, Chennai;
6. Dr Kirtikant Sharma, Indira Gandhi National Centre for Arts, New Delhi;
7. Ms Amba Kulkarni, University of Hyderabad, Hyderabad;
8. Dr Girish Jha, Jawaharlal Nehru University, New Delhi;
9. Dr T Ganesan, French Oriental Institute, Pondicherry;
10. Dr Santosh Kumar Shukla, JNU, New Delhi;
11. Shri Shriramana Sharma, Kanchi Shankara Matham, Tamil Nadu;

12. Shri Prakash Khanore, CDAC, Pune;

13. Ms Neha Gupta, CDAC, Pune;

14. Shri Manoj Jain, Deptt. Of IT, New Delhi.

The various issues with respect to the encoding of the Grantha script in Unicode were discussed considering various proposals submitted to the Unicode Technical Committee. Document L2/10-265R was also referred to in the discussion.

Unification of Grantha with Tamil

The discussion started with the similarity of Tamil and Grantha. Some of the experts suggested that since these two scripts are same only additional characters present in the Grantha should be encoded. The committee discussed the issue and recommended that these are two distinct scripts and also considering the sensitivity involved, these should not be unified. Grantha should be encoded separately.

Characters used for transcription/Characters not used in Sanskrit

Some scholars objected to the presence of characters not used in Sanskrit:

1. Independent vowels and dependent vowel signs for the short vowels E and O,
2. The nukta, and
3. The Dravidian consonants RRA, LLLA and NNNA.

The committee discussed the issue and recommended that though these characters are not used for writing Sanskrit in the Grantha script, these may be useful in transcribing words of other languages like Tamil or English into the Grantha script and hence these should be encoded for Grantha. These characters are:

1130E Letter E,

11312 Letter O,

11329 Letter NNNA,

11331 Letter RRA,

11334 Letter LLLA,

1133C Nukta,

11346 Vowel Sign E and

1134A Vowel Sign O

The committee also recommended to add an annotation to these characters to indicate that they are used for transcription only, as these characters are similar to the Devanagari characters U+090E, U+0912, U+0929, U+0931, U+0934, U+093C, U+0946 and U+094A.

Chandrabindu, Anusvara and Visarga

Though these were not included in the original Gol proposal L2/10-048, it was verified that these are used in the Grantha script and hence should be recommended for encoding.

Two forms of the vowel sign for /au/

The necessity of two forms of vowel sign AU (proposed as U+1134C and U+11357) was discussed. It was pointed that this was as in Malayalam – there is usage of both kinds of writing the dependent vowel. The two-part sign is the “old” form and the single-part sign is the “new” form. Since Malayalam has both forms and is a descendant of Grantha, a correlation can be claimed resulting in the validity of both forms for Grantha. The “new” form is the one used in most contemporary printings. Thus these characters are to be separately encoded. Suitable annotation needs to be added for these two vowel signs.

Two forms of the vowel signs for Vocalic L and LL

At the meeting, Gol inquired about the characters seen in L2/09-372:

11362 Grantha Vowel Sign Vocalic L

1137E Grantha Vowel Sign Vocalic L Old

11363 Grantha Vowel Sign Vocalic LL

1137F Grantha Vowel Sign Vocalic LL Old

It was clarified that the vowel signs for Vocalic L/LL in Grantha may be written either to the side of or below the base consonant. While the form written to the side is the default in modern use, archaic use of the sub-base form is also attested. There is of course no difference in meaning as this is just a different way of writing the same vowel. All other Indic scripts have the vowel signs for Vocalic L/LL placed below the base. The Sanskrit scholars including those working with manuscripts all affirmed this fact. It was further discussed whether this variation can be handled at the font level. Technical experts opined that since the properties of these two forms (spacing against non-spacing) would be different there may be need to encode them separately. The committee recommended the separate encoding of these forms as per judgment of the UTC if the technical reasons involved are found sufficient.

Ligating Virama & Chillu Marker

The experts discussed the need of a ligating virama. It was pointed out that there are different ways of writing vowelless consonants in Grantha. However the experts noted that these vowelless forms are all semantically equivalent and that they should be handled at the font level, not at the encoding level. The committee unanimously recommended that there is no need to separately encode a ligating virama. Since the representations are truly equivalent and there is truly no semantic difference

between the various ways of writing vowelless consonants, the rendering is a freely variable choice that can be worked out in font software.

Gol clearly stated that the chillu-s which have semantic difference are something peculiar to Malayalam, and their atomic encoding for Malayalam was a special case not applicable to other scripts like Grantha. Therefore no ligating virama or chillu marker or any other atomic chillu-s are to be encoded for Grantha.

Since the experts recommended that there is no need to encode any chillus, chillu marker, or any special means of indicating the presence or absence of various vowelless forms, the UTC is requested to remove the ligating virama character from the consensus document L2/10-256R and not to encode any such character now or in the future.

Length Mark

The experts opined that a length mark (U+11355) should not be encoded for Grantha since it is not required even in Telugu (U+0C55) and Kannada (U+0CD5), where a length mark is already encoded.

Grantha OM

It was discussed whether Grantha has any separate symbol for OM. However all the expert scholars agreed that there is no evidence for any such symbol. In Grantha OM is just written as O + anusvara or O + M. So no Grantha OM should be encoded.

Dandas

It was noted that the current status is not to encode separate danda-s for Grantha. Gol pointed out that recently the UTC has been encoding separate danda-s for all the recent Indic scripts to be encoded. However, Gol said that to keep in with the encoding of the major Indic scripts, the status quo as is now can be left to stand. The committee recommended that there is no need to separately encode danda-s for Grantha. The danda-s from the Devanagari block should be used.

Pluta Mark

The experts discussed the usage of this character and agreed that this character is an archaic character and that the contemporary usage is to write the length of a vowel in matra-s within parantheses. The committee recommended to encode this character with an appropriate annotation like “Archaic use only”.

Numerals & Fractions

The fact that Tamil and Grantha share the same set of digits, numbers and fractions was accepted by all without debate. Digits 0-9 and numbers 10, 100 and 1000 are encoded in the Tamil block and can used from the Tamil block. Experts affirmed that there is no need to separately encode Grantha numerals and the Tamil numerals (U+0BE6 to U+0BF2) should be used for Grantha.

Similarly for Grantha fractions, Tamil fractions are to be used. It is understood that the Tamil fractions need to be encoded in Unicode. A separate proposal may be submitted to the Unicode Consortium in consultation with other stakeholders i.e. Govt of Tamil Nadu, INFITT and ICTA Sri Lanka.

Grantha-specific Vedic characters:

The committee discussed the issue of encoding Vedic characters for Grantha. It was agreed that the Vedic characters common to Indian scripts will not be permitted to be duplicated for script-specific blocks but added to the generic Vedic Extensions block. Those characters which are script-specific will certainly have to be encoded in the script's block, or, in the absence of space, in an Extended block as has been done for Devanagari.

Gol expressed a doubt as to whether the Vedic characters for Grantha should be encoded right away, considering the amount of confusion that was prevalent during the Devanagari/generic Vedic encoding for Unicode 5.2. However, the experts pointed out that the user base of Grantha is much smaller than that of Devanagari and there are far fewer publications comparatively. So there are far less variants. Thus there will be no problem in encoding these characters now. It was also informed that the relevant Vedic scholars have been consulted and shown that sufficient attestations for the use of the characters have been provided.

After deliberations, the committee recommended that Vedic characters for Grantha with sufficient attestation may be encoded right now.

Among these the Sama Vedic svara markers proposed for 11366-1136C, 11370-11374 were discussed. These characters are analogous to the Devanagari Sama Vedic characters from the Devanagari Extended block and should be encoded.

Vedic anusvara-s

The characters Vedic Anusvara (1135E) and Vedic Double Anusvara (1135F) were discussed. Dr R K Shastri noted that the double anusvara is also sometimes seen in Devanagari texts. (Apart from publishing Vedic texts in Grantha from his own Heritage India Trust he has also edited Devanagari Vedic texts published by others.) Gol noted that at the time of encoding of the various Vedic characters for Unicode 5.2, very many variants of the anusvara-s were considered, of which some were encoded after much discussion. Gol recommended that if these anusvara forms are seen particularly in Grantha then it is appropriate to encode them for Grantha.

Dr Kirtikant Sharma and Gol remarked that the glyph for the Vedic Double Anusvara as seen in L2/10-265R was somewhat cluttered with the lines too close together. The committee recommended that these Vedic Anusvara characters should be encoded with improvement in the glyph of the Double Anusvara.

Recommendations to the UTC

The experts opined that Gol may forward these recommendations to the UTC so that UTC may finalize the Grantha encoding in the Unicode Standard. In summary the various characters and signs recommended by the committee of experts is as follows with suitable remarks from the meeting.

Grantha characters recommended for encoding

S.No.	Code point and name	Glyph	Remark
VARIOUS CHARACTERS			
1	11301 GRANTHA SIGN CANDRABINDU	◌̣	Gol did not propose the encoding of these signs earlier. After discussion, experts recommended encoding these signs for Grantha. Consensus.
2	11302 GRANTHA SIGN ANUSVARA	◌̣̣	
3	11303 GRANTHA SIGN VISARGA	◌̣̣̣	
INDEPENDENT VOWELS			
4	11305 GRANTHA LETTER A	𑌀	Consensus.
5	11306 GRANTHA LETTER AA	𑌁	Consensus.
6	11307 GRANTHA LETTER I	𑌂	Consensus.
7	11308 GRANTHA LETTER II	𑌃	Consensus.
8	11309 GRANTHA LETTER U	𑌄	Consensus.
9	1130A GRANTHA LETTER UU	𑌅	Consensus.
10	1130B GRANTHA LETTER VOCALIC R	𑌆	Consensus.
11	1130C GRANTHA LETTER VOCALIC L	𑌇	Consensus.
12	1130E GRANTHA LETTER E	𑌈	The committee recommends that though this character is not used for writing Sanskrit in Grantha, it may be used

S.No.	Code point and name	Glyph	Remark
			for transcription of words of other languages like Tamil and English into Grantha and hence should be encoded. Recommended annotation "Used for transcription only". Consensus.
13	1130F GRANTHA LETTER EE	ഈ	Gol earlier proposed this character under the name LETTER E. Name now changed to match pattern of other South Indian scripts. Consensus.
14	11310 GRANTHA LETTER AI	ഈ	Consensus.
15	11312 GRANTHA LETTER O	ഐ	The committee recommends that though this character is not used for writing Sanskrit in Grantha, it may be used for transcription of words of other languages like Tamil and English into Grantha and hence should be encoded. Recommended annotation "Used for transcription only". Consensus.
16	11313 GRANTHA LETTER OO	ഐ	Gol earlier proposed this character under the name LETTER O. Name now changed to match pattern of other South Indian scripts. Consensus.
17	11314 GRANTHA LETTER AU	ഐ	Consensus.
CONSONANTS			
18	11315 GRANTHA LETTER KA	ക	Consensus.
19	11316 GRANTHA LETTER KHA	ഖ	Consensus.

S.No.	Code point and name	Glyph	Remark
20	11317 GRANTHA LETTER GA	𑌧	Consensus.
21	11318 GRANTHA LETTER GHA	𑌨	Consensus.
22	11319 GRANTHA LETTER NGA	𑌩	Consensus.
23	1131A GRANTHA LETTER CA	𑌪	Consensus.
24	1131B GRANTHA LETTER CHA	𑌫	Consensus.
25	1131C GRANTHA LETTER JA	𑌬	Consensus.
26	1131D GRANTHA LETTER JHA	𑌭	Consensus.
27	1131E GRANTHA LETTER NYA	𑌮	Consensus.
28	1131F GRANTHA LETTER TTA	𑌯	Consensus.
29	11320 GRANTHA LETTER TTHA	𑌰	Consensus.
30	11321 GRANTHA LETTER DDA	𑌱	Consensus.
31	11322 GRANTHA LETTER DDHA	𑌲	Consensus.
32	11323 GRANTHA LETTER NNA	𑌳	Consensus.
33	11324 GRANTHA LETTER TA	𑌴	Consensus.
34	11325 GRANTHA LETTER THA	𑌵	Consensus.
35	11326 GRANTHA LETTER DA	𑌶	Consensus.

S.No.	Code point and name	Glyph	Remark
36	11327 GRANTHA LETTER DHA	𑌛	Consensus.
37	11328 GRANTHA LETTER NA	𑌜	Consensus.
38	11329 GRANTHA LETTER NNNA	𑌝	The committee recommends that though this character is not used for writing Sanskrit in Grantha, it may be used for transcription of words of other languages like Tamil and English into Grantha and hence should be encoded. Recommended annotation "Used for transcription only". Consensus.
39	1132A GRANTHA LETTER PA	𑌞	Consensus.
40	1132B GRANTHA LETTER PHA	𑌟	Consensus.
41	1132C GRANTHA LETTER BA	𑌠	Consensus.
42	1132D GRANTHA LETTER BHA	𑌡	Consensus.
43	1132E GRANTHA LETTER MA	𑌢	Consensus.
44	1132F GRANTHA LETTER YA	𑌣	Consensus.
45	11330 GRANTHA LETTER RA	𑌤	Consensus.
46	11331 GRANTHA LETTER RRA	𑌥	The committee recommends that though this character is not used for writing Sanskrit in Grantha, it may be used for transcription of words of other languages like Tamil and English into Grantha and hence should be encoded. Recommended annotation

S.No.	Code point and name	Glyph	Remark
			"Used for transcription only". Consensus.
47	11332 GRANTHA LETTER LA	Ⓦ	Consensus.
48	11333 GRANTHA LETTER LLA	᳚	Consensus.
49	11334 GRANTHA LETTER LLLA	᳛	The committee recommends that though this character is not used for writing Sanskrit in Grantha, it may be used for transcription of words of other languages like Tamil and English into Grantha and hence should be encoded. Recommended annotation "Used for transcription only". Consensus.
50	11335 GRANTHA LETTER VA	᳜	Consensus.
51	11336 GRANTHA LETTER SHA	᳝	Consensus.
52	11337 GRANTHA LETTER SSA	᳞	Consensus.
53	11338 GRANTHA LETTER SA	᳟	Consensus.
54	11339 GRANTHA LETTER HA	᳠	Consensus.
VARIOUS SIGNS			
55	1133C GRANTHA SIGN NUKTA	᳡	The committee recommends that though this character is not used for writing Sanskrit in Grantha, it may be used for transcription of words of other languages like Tamil and English into Grantha and hence should be encoded. Recommended annotation "Used for transcription only". Consensus.

S.No.	Code point and name	Glyph	Remark
56	1133D GRANTHA SIGN AVAGRAHA	𑌛	Consensus.
DEPENDENT VOWEL SIGNS			
57	1133E GRANTHA VOWEL SIGN AA	𑌛𑌃	Consensus.
58	1133F GRANTHA VOWEL SIGN I	𑌛𑌄	Consensus.
59	11340 GRANTHA VOWEL SIGN II	𑌛𑌅	Consensus.
60	11341 GRANTHA VOWEL SIGN U	𑌛𑌆	Consensus.
61	11342 GRANTHA VOWEL SIGN UU	𑌛𑌇	Consensus.
62	11343 GRANTHA VOWEL SIGN VOCALIC R	𑌛𑌈	Consensus.
63	11344 GRANTHA VOWEL SIGN VOCALIC RR	𑌛𑌉	Consensus.
64	11346 GRANTHA VOWEL SIGN E	𑌛𑌊	The committee recommends that though this character is not used for writing Sanskrit in Grantha, it may be used for transcription of words of other languages like Tamil and English into Grantha and hence should be encoded. Recommended annotation "Used for transcription only". Consensus.
65	11347 GRANTHA VOWEL SIGN EE	𑌛𑌋	Gol earlier proposed this character under the name VOWEL SIGN E. Name now changed to match pattern of other South Indian scripts. Consensus.
66	11348 GRANTHA VOWEL SIGN AI	𑌛𑌌	Consensus.

S.No.	Code point and name	Glyph	Remark
67	1134A GRANTHA VOWEL SIGN O	ᱠ	The committee recommends that though this character is not used for writing Sanskrit in Grantha, it may be used for transcription of words of other languages like Tamil and English into Grantha and hence should be encoded. Recommended annotation "Used for transcription only". Consensus.
68	1134B GRANTHA VOWEL SIGN OO	ᱡ	GoI earlier proposed this character under the name VOWEL SIGN O. Name now changed to match pattern of other South Indian scripts. Consensus.
69	1134C GRANTHA VOWEL SIGN AU	ᱢ	The necessity of two forms of the vowel sign for /au/ for was discussed. It was noted that this two-part form is somewhat archaic. Both forms are to be encoded with suitable annotations. Consensus.
VIRAMA			
70	1134D GRANTHA SIGN VIRAMA	ᱣ	Consensus.
DEPENDENT VOWEL SIGN			
71	11357 GRANTHA AU LENGTH MARK	ᱤ	The necessity of two forms of the vowel sign for /au/ for was discussed. It was noted that this one-part form is the modern one. Both forms are to be encoded with suitable annotations. Consensus.
OTHER CHARACTER			
72	1135D GRANTHA SIGN PLUTA	ᱥ	It was noted that this is not in current use but is to be found in manuscripts and should be encoded. Recommended

S.No.	Code point and name	Glyph	Remark
			annotation "Archaic use only". It is also noted that a separate proposal L2/10-331 has been submitted for this. Consensus.
VEDIC CHARACTERS			
73	1135E GRANTHA LETTER VEDIC ANUSVARA	ॐ	Consensus.
74	1135F GRANTHA LETTER VEDIC DOUBLE ANUSVARA	ॐ	Glyph clarity to be improved compared to L2/10-265R. See glyph given here. Consensus.
ADDITIONAL VOWELS			
75	11360 GRANTHA LETTER VOCALIC RR	ꣳ	Consensus.
76	11361 GRANTHA LETTER VOCALIC LL	ꣴ	Consensus.
77	11362 GRANTHA VOWEL SIGN VOCALIC L	ꣵ	The existence of alternative forms of these vowel signs was discussed. These forms are the ones in modern use and should be encoded.
78	11363 GRANTHA VOWEL SIGN VOCALIC LL	ꣶ	For the alternative forms see end of this list. Consensus.
SAMA VEDIC SVARA MARKERS			
79	11366 COMBINING GRANTHA DIGIT ZERO	◌ ◌	These Sama Vedic markers are analogous to Devanagari characters from the Devanagari Extended block and should be encoded. Consensus.
80	11367 COMBINING GRANTHA DIGIT ONE	ꣷ ◌	
81	11368 COMBINING GRANTHA DIGIT TWO	꣸ ◌	
82	11369 COMBINING GRANTHA DIGIT THREE	꣹ ◌	
83	1136A COMBINING GRANTHA DIGIT FOUR	꣺ ◌	

S.No.	Code point and name	Glyph	Remark
84	1136B COMBINING GRANTHA DIGIT FIVE	𑌵 ◌̣	
85	1136C COMBINING GRANTHA DIGIT SIX	𑌶 ◌̣	
86	11370 COMBINING GRANTHA LETTER A	𑌇 ◌̣	<p>These Sama Vedic markers are analogous to Devanagari characters from the Devanagari Extended block and should be encoded.</p> <p>Consensus.</p>
87	11371 COMBINING GRANTHA LETTER KA	𑌈 ◌̣	
88	11372 COMBINING GRANTHA LETTER NA	𑌉 ◌̣	
89	11373 COMBINING GRANTHA LETTER VI	𑌊 ◌̣	
90	11374 COMBINING GRANTHA LETTER PA	𑌋 ◌̣	
ADDITIONAL VOWEL SIGNS			
91	1137E GRANTHA VOWEL SIGN SUB-BASE VOCALIC L	𑌌 ◌̣	<p>The existence of alternative forms of these vowel signs was discussed. These forms are mostly in archaic use. It may be necessary to encode these archaic forms separately on technical grounds. It is noted that a separate document L2/10-341 has been submitted regarding this.</p> <p>Encoding recommended if technical grounds are found sufficient.</p>
92	1137E GRANTHA VOWEL SIGN SUB-BASE VOCALIC LL	𑌍 ◌̣	

Characters recommended for not encoding

	GRANTHA SIGN LIGATING VIRAMA		Experts recommended that the intended purpose of ligating virama is to be taken care of by appropriate font tables etc. This character should <u>not</u> be encoded.
	GRANTHA SIGN CHILLU MARKER		Gol and experts recommended that the chillu-s are something peculiar to Malayalam and their atomic encoding for Malayalam was a special case not applicable to other scripts like Grantha. Therefore <u>no</u> chillu marker or any other atomic chillu-s are to be encoded for Grantha.
	GRANTHA OM		Experts opined that there is no separate sign for Grantha OM, hence it should not be encoded.
	GRANTHA LENGTH MARK		Not to be encoded as it is not used distinctly in the script.
	GRANTHA DANDA GRANTHA DOUBLE DANDA		Experts recommended using the danda-s from the Devanagari block.
	GRANTHA NUMERALS AND FRACTIONS		
	GRANTHA DIGIT ZERO GRANTHA DIGIT ONE GRANTHA DIGIT TWO GRANTHA DIGIT THREE GRANTHA DIGIT FOUR GRANTHA DIGIT FIVE GRANTHA DIGIT SIX GRANTHA DIGIT SEVEN GRANTHA DIGIT EIGHT GRANTHA DIGIT NINE GRANTHA NUMBER TEN GRANTHA NUMBER ONE HUNDRED GRANTHA NUMBER ONE THOUSAND		The fact that Tamil and Grantha share the same set of digits, numbers and fractions was accepted by all without debate. Digits 0-9 and numbers 10, 100 and 1000 are encoded in the Tamil block and can be used from the Tamil block. Experts affirmed that there is no need to separately encode Grantha numerals. Tamil numerals (U+0BE6 to U+0BF2) should be used for Grantha.

	<p>GRANTHA FRACTION 80/320 GRANTHA FRACTION 160/320 GRANTHA FRACTION 240/320 GRANTHA FRACTION 1/320 GRANTHA FRACTION 2/320 GRANTHA FRACTION 4/320 GRANTHA FRACTION 8/320 GRANTHA FRACTION 12/320 GRANTHA FRACTION 16/320 GRANTHA FRACTION 20/320 GRANTHA FRACTION 32/320 GRANTHA FRACTION 40/320 GRANTHA FRACTION 48/320 GRANTHA FRACTION 60/320 GRANTHA FRACTION 64/320</p>		<p>The fact that Tamil and Grantha share the same set of digits, numbers and fractions was accepted by all without debate.</p> <p>Therefore, for Grantha fractions, Tamil fractions may be used. It is understood that the Tamil fractions need to be encoded in Unicode. A separate proposal may be submitted to the Unicode Consortium in consultation with other stakeholders.</p>
--	--	--	---

The Department of Information Technology, Government of India will be submitting a separate complete proposal with proposal summary form, code chart and Unicode character properties for all the characters recommended here. It also recognizes that for the pluta sign and archaic vowel signs of vocalic L/LL, separate proposals have been already submitted. The Government of India recommends the UTC to further the encoding of the Grantha script based on these proposal documents.

Note: The Grantha font for use with the proposal and code chart was very kindly provided by Mr Elmar Kniprath.

-0-