

Proposal to encode 09FC BENGALI LETTER VEDIC ANUSVARA

Shriramana Sharma, jamadagni-at-gmail-dot-com, India

2015-Jul-31

L2/15-161R

This is a proposal to encode one character in the Bengali block:



09FC BENGALI LETTER VEDIC ANUSVARA

This character is required for representation of Vedic texts – especially of the Taittiriya school of the Krishna Yajur Veda and the Kauthuma school of the Sama Veda – in the Bengali script. It denotes a Vedic anusvara, and is found in contrast to the regular anusvara. Attestations from old publications showing the contrast are provided in this document.

Parallel characters in other scripts

A8F2 ॡ DEVANAGARI SIGN SPACING CANDRABINDU and A8F3 ॢ DEVANAGARI SIGN CANDRABINDU VIRAMA were contrastively attested for the Sama Vedic and Yajur Vedic usages respectively in L2/07-343 p 25. 1135E GRANTHA LETTER VEDIC ANUSVARA ॣ was attested in L2/09-372 p 12 for both these usages. 0C80 KANNADA SIGN SPACING CANDRABINDU । was proposed for Badaga language orthography by L2/14-153 and attestation for its usage for the Sama Veda has been provided in L2/15-158. Since all these characters have been encoded per-script, it is proposed to encode the parallel character for Bengali separately.

To be noted is that the Devanagari, Grantha, Kannada characters used for the Sama Veda do *not* have a horizontal downward stroke below. However in Bengali, it is only the form *with* the stroke below that is attested for both the Sama Veda and Yajur Veda. Similar linguistic contexts from the Rig Veda and Atharva Veda use the non-spacing combining character 0981 BENGALI SIGN CANDRABINDU ॥.

Justification for separate encoding

The proposed character resembles 09FA BENGALI ISSHAR १ combined with 09CD BENGALI SIGN VIRAMA ॡ. However, the VEDIC ANUSVARA has no linguistic connection with the ISSHAR.

It is unclear why the BENGALI ISSHAR sign which denotes “heavenly” beings resembles a candrabindu. The representative form of the cognate character 0B70 ORIYA ISSHAR ॡ

resembles a “tick”-mark (“check”) and not a candrabindu. It thus appears the ISSHAR is not tied down to a rounded shape whereas the VEDIC ANUSVARA is only attested with this shape.

Technically speaking also, the ISSHAR is GC=So and having it combine with the VIRAMA is likely to cause unnecessary complications in implementation due to word-break properties etc. Thus it is recommended to encode the VEDIC ANUSVARA separately.

Character name

Considering the related and already encoded characters from Devanagari, Grantha and Kannada, there are two possibilities for the name: LETTER VEDIC ANUSVARA (from Grantha) and SIGN CANDRABINDU VIRAMA (from Devanagari), the one per the linguistic sense and the other per the orthographic form. Of these, we choose the former.

The latter name is appropriate for the Devanagari character to differentiate from the plain SPACING CANDRABINDU since in that script there is a contrast between the usage and absence of the downward stroke. However in Bengali, a future publication may well choose a glyph without the downward stroke for simplicity since there is no attested contrast. It is preferable that the encoding should not pre-empt such a stylistic freedom of the users by naming the character otherwise. Hence the recommended name is LETTER VEDIC ANUSVARA.

Unicode Character Properties

09FC;BENGALI LETTER VEDIC ANUSVARA;Lo;0;L;;;;;N;;;;;

The code point 09FC is chosen to be close to 09FA BENGALI ISSHAR [✓] to facilitate font-makers to become aware of the need to synchronize the appearances of these two characters which are very similar.

Other properties such as line-breaking, Indic syllabic category etc may be taken as that of 1135E GRANTHA LETTER VEDIC ANUSVARA. For collation, the character (and other similar characters for Vedic usage) should be treated equivalent with the anusvara, since it only occurs in place of the latter in specific contexts.

Attestations

(all from texts edited by Durgadas Lahiri and published in Kolkota)

Taittiriya school of the Krishna Yajur Veda

Vol. 1, pub. 1926, pages 1, 24, 271-273

प्यायध्वमग्नि॑या॒ दे॒व॒भ॒ग॒मूर्ज॑स॒वतीः॒ पर॑स॒वतीः॒ प्र॒जा॒व॒ती॒र-
न॒मी॒वा॒ अ॒य॒क्ष्मा॒ मा॒ वः॒ स्तेन॑ इ॒श॒त॒ मा॒श्व॑सा॒से।

āpyāyadhvamagnīyā devabhāgamūrjasvatīḥ payasvatīḥ prajāvatīra-
namībā ayakṣmā mā baḥ stena īśata māśvasāśso

(१) य॒ज्ज॒स्र॒ घो॒ष॒द॒सि॒ । (२) प्र॒त्यू॒ष्टा॒ र॒क्षः॒ प्र॒त्यू॒ष्टा॒ अ॒रो॒त॒यः॒ ।
(३) प्रे॒य॒मा॒ग॒द्धि॒षा॒णा॒ ब॒र्हि॒र॒च्छ॒ म॒नु॒ना॒ कृ॒ता॒ स्व॒ध॒या॒ वि॒त॒ष्टा॒ त॒ आ॒
ब॒ह॒न्ति॒ क॒व॒यः॒ पु॒र॒स्ता॒द्दे॒वे॒भ्यो॒ जु॒ष्ट॒मि॒ह॒ ब॒र्हि॒र॒सा॒दे॒ ।
(४) दे॒वा॒नां॒ परि॒षू॒त॒म॒सि॒ व॒र्ष॒बृ॒द्ध॒म॒सि॒ ।

yajñasya ghoṣadasi | pratyuṣṭaṁ rakṣaḥ pratyuṣṭā arotayaḥ |
preyamaḡāddhiṣaṇā barhiraccha manunā kṛtā svadhayā vitaṣṭā ta ā
bahanti kabayaḥ purastāddevebhyo juṣṭamiha barhirāsade
debānām pariṣūtamasi barṣabṛddhamasi |

বাজস্য মা প্রসবেনোদ্গ্রাভেণোদগ্রভীৎ অথা সপত্না ইন্দ্রো
 মে নিগ্রাভেণাধরাৎ অকঃ । উদ্গ্রাভং চ নিগ্রাভং

bājasya mā prasabenodgrābheṇodagrabhīt | athā sapatnāṁ indro
 me nigrābheṇādharāṁ akaḥ | udgrābham ca nigrābham

যং পরিধিৎ পর্যাদত্থা অগ্নে দেব পণিভিব্বীয়মাণঃ । তং ত
 এতমনু জোষং ভরামি নেদেষ ত্বদপচেতয়াতৈ
 যজস্য পাথ উপ সমিতুৎ ।

সৎ স্রাবভাগাঃ স্বেষা বৃহন্তঃ প্রস্তরেষ্ঠা বহিষদধ দেবা ইমাং

yaṁ paridhiṁ paryadhatthā agne deba paṇibhirbīyamāṇaḥ | taṁ ta
 etamanu joṣaṁ bharāmi nedeṣa tvadapacetayātai
 yajñasya pātha upa samitaṁ |

saṁsrābhāgāḥ stheṣā bṛhantaḥ prastareṣṭhā barhiṣadaśca debā imāṁ

পাছি | দুৱদায়ে | পাছি | দুশ্চৰিতাদবিষং | নাঃ | পিতৃং |
 কৃণু | সুসদা | যোনিম্ | স্বাহা |

pāhi duradmanyai pāhi duścaritādaviṣam naḥ pituṃ
 kṛṇu suśadā yonim̐ svāhā |

Kauthuma school of the Sama Veda

Vol. 1, pub. 1922, Agneya, pages 15 and 35; Vol. 5, pub. 1927, Mahānāmni, page 7

প্ৰেষ্ঠং | বো | অতিথিম্ | স্তুষে | মিত্ৰমিবা | প্ৰিয়ং |
 অগ্নে | রথং | ন | বেদ্যং | ৫ ||

preṣṭham̐ bo atithim̐ stuṣe mitramiba priyam |
 agne ratham̐ na bedyam̐ ||

দূতং | বো | বিশ্ববেদসম্ | হব্যাবাহমমর্ত্যং |
 যজিষ্ঠমৃঞ্জসে | গিরঃ || ২ ||

dūtam̐ bo biśvabedasaṃ habyaabāhamamartyam̐ |
 yajiṣṭhamṛñjase giraḥ ||

ॐ २ ० २ ० २ ० १ २ ० २ ० २ ० १ २
विदा राये सुवीर्यास्तुवो बाजानास्पतिवर्षा^७ अनु ।

३ २ ० २ ० १ २ २ ० १ २
म^७हिष्ठ बज्रिन् ऋजसेयः शविष्ठ शूराणाम् ॥ ४ ॥
* * *

पङ्कमं गाम ।

१ २ २ ० १ २ ० २ ० १ २
यो म^७हिष्ठो मघोनाम् अ^७शुः न शोचिः ।

bidā rāye subīryambhavo bājānāmpatirbbaśāṁ anu |
maṁhiṣṭha bajrin ṛñjaseyaḥ śabiṣṭha śūrāṇām ||
yo maṁhiṣṭho maghonām aṁśuḥ na śociḥ |

Attestation for non-spacing candrabindu used elsewhere

Shakala school of the Rig Veda, vol. 1, part 1, pub. 1925, page 38

शतं॑ रा॒ज्जे॒ना॒ध॒मा॒नस्य॑ नि॒ष्का॒ञ्च॒त॒मा॒स॒व॒न् प्र॒य॒त॒ान्तं॑ स॒द्य॒ आ॒दि॒म् ।
श॒त॒म् क॒क्षि॒बा॒म् अ॒सुर॑स्य॒ गो॒ना॒म् दि॒वि श्र॒बो ज॒र॒मा॒त॒त॒ान् ॥ २ ॥

śatam rājñonādhamānasya niṣkāñchatamaśvān prayatānt sadya ādam |
śatam kakṣibāṁ asurasya gonām dibi śrabo jaramātātāna ||

Shaunaka school of the Atharva Veda vol. 1, pub. 1910, page 103

त्व॒मग्ने॑ या॒तु॒धा॒ना॒न् उ॒प॒ब॒द्ध॑म् । इ॒हा॒ब॒हा ।

tvamagne yatudhānān upabaddhām ihābaha |

Official Proposal Summary Form

(Based on N4502-F)

A. Administrative

1. Title

Proposal to encode 09FC BENGALI LETTER VEDIC ANUSVARA

2. Requester's name

Shriramana Sharma

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution

4. Submission date

2015-Jul-31

5. Requester's reference (if applicable)

6. Choose one of the following: This is a complete proposal (or) More information will be provided later

This is a complete proposal.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters), Proposed name of script

No

1b. The proposal is for addition of character(s) to an existing block, Name of the existing block

Bengali

2. Number of characters in proposal

1 (one)

3. Proposed category

Category B1, specialized small (for this character, though Bengali itself is “A, contemporary”)

4. Is a repertoire including character names provided?

Yes

4a. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?

Yes

4b. Are the character shapes attached in a legible form suitable for review?

Yes

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?

Shriramana Sharma

b. Identify the party granting a license for use of the font by the editors (include address, e-mail etc.)

Shriramana Sharma

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.

See detailed proposal.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

Yes. The non-R version dated 2015-May-18 was submitted with the proposed codepoint as 09CF. Based on the recommendation from the script ad hoc report L2/15-204 endorsed at the 2015 Jul UTC meeting, the codepoint has now been changed to 09FC. No other changes were made.

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?

Vidvan Gagan Kumar Chatterjee, native speaker/user of the Bengali language/script, scholar of the Kauthuma school of the Sama Veda, currently resident in Varanasi, India.

2c. If YES, available relevant documents

None specifically. The matter was discussed in person and via email/phone.

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Those who wish to digitize Bengali script texts of the Krishna Yajur Veda and Kauthuma Sama Veda.

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

Rare. Old usage is attested.

4b. Reference

See detailed proposal.

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

No.

5b. If YES, where?

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?

Yes

6c. If YES, reference

The character belongs to the Bengali block where sufficient space is available.

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

Only one character is proposed.

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, or could be confused with, an existing character?

The character is similar to A8F3 DEVANAGARI SIGN CANDRABINDU VIRAMA.

10b. If YES, is a rationale for its inclusion provided?

Yes

10c. If YES, reference

Existing practice is to encode Vedic anusvara characters per-script.

11a. Does the proposal include use of combining characters and/or use of composite sequences?

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?

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?

13c. If YES, reference:

-O-O-O-