Proposal to Encode the Ganda Currency Mark for Bengali in ISO/IEC 10646

Anshuman Pandey
University of Michigan
Ann Arbor, Michigan, U.S.A.
pandey@umich.edu

September 25, 2007

Contents

List of Figures

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]

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: **Proposal to Encode the Ganda Currency Mark for Bengali in ISO/IEC 10646**
2. Requester's name: **Anshuman Pandey (pandey@umich.edu)**
3. Requester type (Member Body/Liaison/Individual contribution): **Individual contribution**
4. Submission date: **September 25, 2007**
5. Requester's reference (if applicable): **N/A**
6. Choose one of the following:
   (a) This is a complete proposal: **Yes**
   (b) or, More information will be provided later: **No**

B. Technical - General

1. Choose one of the following:
   (a) This proposal is for a new script (set of characters): **No**
       i. Proposed name of script: **N/A**
   (b) The proposal is for addition of character(s) to an existing block: **Yes**
       i. Name of the existing block: **Bengali**
2. Number of characters in proposal: **1**
3. Proposed category: **A - Contemporary**
4. Is a repertoire including character names provided?: **Yes**
   (a) If Yes, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?:
       Yes
   (b) Are the character shapes attached in a legible form suitable for review?: **Yes**
5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for
   publishing the standard?: **Anshuman Pandey**; **True Type**
   (a) If available now, identify source(s) for the font and indicate the tools used: **The font contains a normalized
       form of the sign as found in printed documents. It was drawn by Anshuman Pandey using Metafont
       and converted to True Type format using FontForge.**
6. References:
   (a) Are references (to other character sets, dictionaries, descriptive texts etc.) provided?: **Yes**
   (b) Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed
       characters attached?: **Yes**
7. Special encoding issues:
   (a) 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)? **No**
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. **Character properties, numeric information, and currency information are included.**

---

[1]Form number: N3102-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)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?: **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.)? **No**
   (a) If Yes, with whom?: **N/A**
       i. If Yes, available relevant documents: **N/A**
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? **Yes**
   (a) Reference: **The sign was used by the Bengali-speaking community.**
4. The context of use for the proposed characters (type of use; common or rare): **Common**
   (a) Reference: **The sign was used to write currency notation in the Bengali script.**
5. Are the proposed characters in current use by the user community?: **No**
   (a) If Yes, where? Reference: **The sign is not used at present.**
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?: **Yes**
   (a) If Yes, is a rationale provided?: **The sign belongs to the Bengali script, which is encoded in the BMP. There is sufficient space in the Bengali block for the inclusion of this sign.**
       i. If Yes, reference: **N/A**
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? **Yes. It should be added to a code-point adjacent to other currency signs in the Bengali block.**
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? **No**
   (a) If Yes, is a rationale for its inclusion provided?: **N/A**
       i. If Yes, reference: **N/A**
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? **No**
   (a) If Yes, is a rationale provided?: **N/A**
       i. If Yes, reference: **N/A**
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? **Yes**
    (a) If Yes, is a rationale for its inclusion provided? **Yes**
        i. If Yes, reference: **Similar in function to other currency marks. See text of proposal for additional details.**
11. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)? **No**
    (a) If Yes, is a rationale for such use provided? **N/A**
        i. If Yes, reference: **N/A**
    (b) Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? **No**
        i. If Yes, reference: **N/A**
12. Does the proposal contain characters with any special properties such as control function or similar semantics? **No**
    (a) If Yes, describe in detail (include attachment if necessary): **N/A**
13. Does the proposal contain any Ideographic compatibility character(s)? **No**
    (a) If Yes, is the equivalent corresponding unified ideographic character(s) identified? **N/A**
        i. If Yes, reference: **N/A**

1   Introduction

This is a proposal to encode the ﹨ BENGALI GANDA MARK as part of the Bengali script in the Basic Multilingual Plane (BMP) of the Universal Character Set (UCS) (ISO/IEC 10646). The intention is to provide a sign used for writing the historical *gaṇḍā* (গণ্ডা) currency unit. Several signs used for the writing of currency and other quantities in Bengali are already encoded in the UCS:

﹨	U+09F2	BENGALI RUPEE MARK
৳	U+09F3	BENGALI RUPEE SIGN
৴	U+09F4	BENGALI CURRENCY NUMERATOR ONE
৵	U+09F5	BENGALI CURRENCY NUMERATOR TWO
৶	U+09F6	BENGALI CURRENCY NUMERATOR THREE
৷	U+09F7	BENGALI CURRENCY NUMERATOR FOUR
৸	U+09F8	BENGALI CURRENCY NUMERATOR ONE LESS THAN THE DENOMINATOR
৹	U+09F9	BENGALI CURRENCY DENOMINATOR SIXTEEN

The encoding of BENGALI GANDA MARK is necessary for accurately and fully reproducing historical numerical notation in the Bengali script and for the representation of the currency mark in digital media.

2   Overview of the Mark

﹨   BENGALI GANDA MARK

Name   The mark is named the BENGALI GANDA MARK. The name describes its function as the unit mark for the *gaṇḍā* currency unit.

Description   The BENGALI GANDA MARK belongs to a system of currency notation used in Bengal and other areas of eastern India. The BENGALI GANDA MARK is one of three distinct signs in the Bengali script used for writing the historical currency units *rupayā* (রূপয়) [or *ṭākā* (টাকা)], *ānā* (আনা), and *gaṇḍā* (গণ্ডা). All three signs appear in written and printed materials. Signs for writing *rupayā* and and *ānā* values are already encoded in the UCS. The sign for the *rupayā* is encoded as ﹨ U+09F2 BENGALI RUPEE MARK and the sign for the *ānā* appears as ৹ U+09F9 BENGALI CURRENCY DENOMINATOR SIXTEEN. The BENGALI GANDA MARK is semantically similar to other currency signs encoded in the UCS, such as the ¢ U+00A2 CENT SIGN, which represent specific currency sub-units. Other systems of currency notation, such as the one used throughout north India, employ a single sign for writing all currency units.[1]

Allocation   It is recommended that the BENGALI GANDA MARK be encoded at the code point U+09FB. The placement is appropriate since the preceding code points (U+09F2–U+09F9) are currency signs, with the exception of U+09FA BENGALI ISSHAR.

Properties   The BENGALI GANDA MARK belongs to the Unicode general category "Symbol, Currency" (Sc). It is a non-combining sign. Similar to other currency signs, it has a bidirectional value of "European Number Terminator" (ET). Its properties in the Unicode Character Database format are:

```
09FB;BENGALI GANDA MARK;Sc;0;ET;;;;;N;;;;;
```

---

[1]See the description of the north Indian currency notation system given in Pandey (2007).

3   Bengali Currency Notation

Like other Indian systems, the Bengali currency system is based on the *rupayā*, Anglicized as "rupee." In Bengal, the *rupayā* is also referred to as *ṭākā*. Historically, the *rupayā* is comprised of a smaller unit called the *ānā*; there are 16 *ānā* in 1 *rupayā*. The *ānā* consists of a smaller unit called the *gaṇḍā*; there are 20 *gaṇḍā* in 1 *ānā*. Each unit has a distinct orthography:

- The *rupayā* is indicated with digits and is marked with ৲ U+09F2 ʙᴇɴɢᴀʟɪ ʀᴜᴘᴇᴇ ᴍᴀʀᴋ. The mark is written after the unit: ৭৲ "7 *rupayā*."

- The *ānā* is written with currency numerators. It is marked with ৹ U+09F9 ʙᴇɴɢᴀʟɪ ᴄᴜʀʀᴇɴᴄʏ ᴅᴇ-ɴᴏᴍɪɴᴀᴛᴏʀ sɪxᴛᴇᴇɴ, which is written after the unit (see Figure 2):

1 *ānā*	৴৹	5 *ānā*	৷৴৹	9 *ānā*	৷৷৴৹	13 *ānā*	৸৴৹
2 *ānā*	৵৹	6 *ānā*	৷৵৹	10 *ānā*	৷৷৵৹	14 *ānā*	৸৵৹
3 *ānā*	৶৹	7 *ānā*	৷৶৹	11 *ānā*	৷৷৶৹	15 *ānā*	৸৶৹
4 *ānā*	৷৹	8 *ānā*	৷৷৹	12 *ānā*	৸৹	1 *rupayā*	১৲

- The *gaṇḍā* is written using digits and is placed after the ʙᴇɴɢᴀʟɪ ɢᴀɴᴅᴀ ᴍᴀʀᴋ (see Figure 1):

1 *gaṇḍā*	৻১	6 *gaṇḍā*	৻৬	11 *gaṇḍā*	৻১১	16 *gaṇḍā*	৻১৬
2 *gaṇḍā*	৻২	7 *gaṇḍā*	৻৭	12 *gaṇḍā*	৻১২	17 *gaṇḍā*	৻১৭
3 *gaṇḍā*	৻৩	8 *gaṇḍā*	৻৮	13 *gaṇḍā*	৻১৩	18 *gaṇḍā*	৻১৮
4 *gaṇḍā*	৻৪	9 *gaṇḍā*	৻৯	14 *gaṇḍā*	৻১৪	19 *gaṇḍā*	৻১৯
5 *gaṇḍā*	৻৫	10 *gaṇḍā*	৻১০	15 *gaṇḍā*	৻১৫	1 *ānā*	৴৹

The *gaṇḍā* is divided into an intermediate unit called the *pāī* (পাঈ). There are 5 *gaṇḍā* in 1 *pāī*. Thus, ৻৫ is equal to 1 *pāī*, ৻১০ to 2 *pāī*, etc., resulting in 4 *pāī* in 1 *ānā*. Despite the name, the Bengali *pāī* differs in value from the north Indian *pāī*. The Bengali *pāī* unit is similar to the north Indian *paisā* (পযসা or পয়সা), also an intermediate unit; there are 4 *paisā* in 1 *ānā*. The *gaṇḍā* unit is closer to the north Indian *pāī*; there are 12 *pāī* in 1 *ānā* in the latter system.[2]

- Historically, there is a unit smaller than the *gaṇḍā* called কড়ি *kaṛi*. There are 4 *kaṛi* in 1 *gaṇḍā*. The *kaṛi* is written using ʙᴇɴɢᴀʟɪ ᴄᴜʀʀᴇɴᴄʏ ɴᴜᴍᴇʀᴀᴛᴏʀ ꜰᴏᴜʀ and ʙᴇɴɢᴀʟɪ ᴄᴜʀʀᴇɴᴄʏ ɴᴜᴍᴇʀᴀᴛᴏʀ ᴏɴᴇ ʟᴇss ᴛʜᴀɴ ᴛʜᴇ ᴅᴇɴᴏᴍɪɴᴀᴛᴏʀ. It is marked by the ʙᴇɴɢᴀʟɪ ᴄᴜʀʀᴇɴᴄʏ ᴅᴇɴᴏᴍɪɴᴀᴛᴏʀ sɪxᴛᴇᴇɴ, which is written before the unit. Thus, 1 *kaṛi* is ৹৷, 2 *kaṛi* is ৹৷৷, and 3 *kaṛi* is ৹৸.

Only one currency mark is used when multiple units are written. This mark is typically the sign for the smallest unit. When *rupayā* and *ānā* values are written together, the Bengali rupee mark is dropped and only ʙᴇɴɢᴀʟɪ ᴄᴜʀʀᴇɴᴄʏ ᴅᴇɴᴏᴍɪɴᴀᴛᴏʀ sɪxᴛᴇᴇɴ is used: "15 *rupayā* and 3 *ānā*" is written as ১৫৶৹, not as ১৫৲৶৹ or as ১৫৲৶. When *rupayā*, *ānā*, and *gaṇḍā* are written together, only the *gaṇḍā* mark is used: "27 *rupayā*, 6 *ānā*, and 5 *gaṇḍā*" is written as ২৭৷৵৻৫. When a value contains *kaṛi*, only ʙᴇɴɢᴀʟɪ ᴄᴜʀʀᴇɴᴄʏ ᴅᴇɴᴏᴍɪɴᴀᴛᴏʀ sɪxᴛᴇᴇɴ is used: "2 *rupayā*, 5 *ānā*, and 3 *kaṛi*" is written as ২৷৴৹৸.

Comparison to Other Systems   The Bengali notation system differs entirely from the system used in Bihar and Uttar Pradesh. The Bengali system uses currency numerators for writing fraction values, while the north Indian system uses additive fraction signs. Also, while the currency units are written with distinctive signs in Bengali, the rupee mark ৲ is used for writing all units in the north Indian system. Thus, "3 *rupayā*" is ৩৲ in Bengali and ३৲ in Devanagari; "15 *ānā*" is ৸৶৹ in Bengali and ⅢӠ৲ in Devanagari. The north Indian

---

[2]Halhed, 1778: 176–177.

pāī is written using fractions and appears after the currency sign, for example, ⌡⸗ "2 *pāī*". The equivalent Bengali unit, the *gaṇḍā*, is written using digits and the BENGALI GANDA MARK, ৻২.

Modern Notation   The use of currency marks and currency numerators in the Bengali script diminished in the latter half of the 20th century when India changed the base of its currency. On April 1, 1957, India introduced a new coinage system called "Naya Paisa," which is based on the decimal system. While the *rupayā* unit was retained, the *ānā* and *gaṇḍā* denominations were replaced with *nayā paisā* (নয়া পয়সা). There are now 100 *paisā* in 1 *rupayā*, instead of the previous 16 *ānā* and 80 *gaṇḍā*. Currency is now written using digits and the rupee mark has been replaced with the Latin 'Rs.' U+20A8 RUPEE SIGN ('Re.' is used for a single rupee). The new rupee sign is written in Indic scripts as the syllable *ru*, which is an abbreviation for *rupayā*; for example, Bengali কৃ৹. The Bengali script possesses the ৳ U+09F3 BENGALI RUPEE SIGN, which is also used for writing the modern rupee. The Latin and Bengali rupee signs are written before the currency. However, there are several ways of separating currency units, eg. using a solidus, dash, period, and other Latin punctuation.

4   Proposal History

This proposal was submitted to the Unicode Technical Committee on May 21, 2007 under the title "Proposal to Encode the Ganda Currency Mark for Bengali in the BMP of the UCS" (L2/07-192). The UTC accepted the proposal

This proposal is a revision of the document submitted to the Unicode Technical Committee (UTC), titled "Proposal to Encode the Ganda Currency Mark for Bengali in the BMP of the UCS" (L2/07-192; ISO/IEC JTC1/SC2/WG2 N3311). The UTC accepted L2/07-192 on August 9, 2007. The character was tentatively allocated at U+09FB, as proposed.[3]

5   References

Beri, D. C. 19–?. हिन्दी बंगला टीचर [*Hindī Baṃgalā Ṭīcara* = Hindi Bengali Teacher]. Calcutta: Hindi Pracharak Pustakalaya.

Everson, Michael. 2007. "Result of Repertoire Review for PDAMx of ISO/IEC 10646:2003 and future amendments." ISO/IEC JTC1/SC2/WG2/N3321 and L2/07-286. August 26, 2007. `http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3316.pdf`.

Grierson, G.A. 1903. *The Linguistic Survey of India*. Vol. V. Indo-Aryan Family. Eastern Group. Part I. Specimens of the Bengali and Assamese languages. Calcutta: Office of the Superintendent of Government Printing, India.

Halhed, Nathaniel Brassey. 1778. *A Grammar of the Bengal Language*. Hoogly, Bengal.

Hudson, D. F. 1965. *Teach Yourself Bengali*. London: The English Universities Press.

Pandey, Anshuman. 2007. "Proposal to Encode North Indic Number Forms in ISO/IEC 10646." L2/07-139R. September 25, 2007. `http://www-personal.umich.edu/~pandey/northindicnumberforms.pdf`

United States National Body, International Organization for Standardization. 2007. "Proposed additions to ISO/IEC 10646:2003." ISO/IEC JTC1/SC2/WG2/N3321. September 7, 2007. `http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3321.pdf`.

---

[3]Everson, 2007; United States National Body, ISO, 2007: 2.

रुपया आना पाई लिखने का तरीका ।

Figure 1: The method of transliterating Bengali currency notation to the north Indian system (from Beri, 19–?: 21). Note the use of ꠂ BENGALI GANDA MARK for writing *gaṇḍā* values. In north Indian notation, the *gaṇḍā* is written in the manner used for *pāī*. The specimen has two typographical errors. First, Bengali /ৎ is incorrectly transliterated as ¬)) (rows 3 and 4, column 1); the correct form is ¬)।, as in the transliteration of /১০ and /১৫. Second, Bengali ১২॥ is incorrectly transliterated as ৽)=); the correct form is ৽)=, without the second NORTH INDIAN RUPEE MARK, as in the transliteration of ১৫, ১১০, etc.

'The leading feature in Indian arithmetic being the division by four, the signs for fractions are adapted thereto. The rupee is divided into 4 × 4 = 16 parts, called *ānā* which are thus designated (units of all kinds are also thus divided) :—

1 ānā or 1/16	/०	5 ānās	।/०	9 ānās	॥/०	13 ānās	৸/०
2 ānās	৶/०	6 ānās	।৶/०	10 ānās	॥৶/०	14 ānās	৸৶/०
3 ānās	৷/०	7 ānās	।৷/०	11 ānās	॥৷/०	15 ānās	৸৷/०
4 ānās or 1/4	।०	8 ānās or 1/2	॥०	12 ānās or 3/4	৸०		

Figure 2: The method of writing fractions of the *ānā* currency unit (from Grierson, 1903b: 29).

For writing money the ordinary numerals are used with hasanta for full rupees and smaller amounts are expressed by the following symbols:—

/०—1 anna   ।/०—5 annas   ॥/०—9 annas   ৸/०—13 annas
৶०—2 annas   ।৶०—6 annas   ॥৶०—10 annas   ৸৶०—14 annas
৷०—3 annas   ।৷०—7 annas   ॥৷०—11 annas   ৸৷०—15 annas
।०—4 annas   ॥०—8 annas   ৸०—12 annas   ১\—1 rupee

Figure 3: Bengali currency notation showing use of Bengali rupee sign (from Hudson, 1965: 85). The description states that the *hasanta* (BENGALI SIGN VIRAMA) is used to write rupee values. Such substitution occurs in print when the glyph for BENGALI RUPEE MARK is absent from a font.