

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

October 8, 2007

Contents

Proposal Summary Form	i
1 Introduction	1
2 Overview of the Mark	1
3 Bengali Currency Notation	2
4 Proposal History	3
5 References	3

List of Figures

1	The conversion of Bengali currency notation to the north Indic system	4
2	The orthography of currency in Bengali.	4
3	Fractions of the <i>ānā</i> currency unit.	5
4	Bengali currency notation showing use of Bengali rupee sign	5

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹

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: **October 8, 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 character 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.**

¹Form 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?: **Yes. This proposal is a revision of “Proposal to Encode the Ganda Currency Mark for Bengali in the BMP of the UCS” (ISO/IEC JTC1/SC2/WG2 N3311 L2/07-192).**
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 character 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 character 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 character 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 character belongs to the Bengali script, which is encoded in the BMP. There is sufficient space in the Bengali block for the inclusion of this character.**
 - i. If Yes, reference: **N/A**
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? **Yes. The character should be encoded at 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: **The character is 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? **N/A**
 - 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**

3 Bengali Currency Notation

Similar to other Indic currency systems, the Bengali 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*, anglicized as “anna”; there are 16 *ānā* in 1 *rupayā*. The *ānā* itself 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 BENGALI RUPEE MARK. The mark is written after the unit: ১ “1 *rupayā*.”
- The *ānā* is written with ‘currency numerators,’ or fraction signs. It is marked with U+09F9 BENGALI CURRENCY DENOMINATOR SIXTEEN, which is written after the unit (see Figure 3):

1 <i>ānā</i>	১	5 <i>ānā</i>	৫	9 <i>ānā</i>	৯	13 <i>ānā</i>	১৩
2 <i>ānā</i>	২	6 <i>ānā</i>	৬	10 <i>ānā</i>	১০	14 <i>ānā</i>	১৪
3 <i>ānā</i>	৩	7 <i>ānā</i>	৭	11 <i>ānā</i>	১১	15 <i>ānā</i>	১৫
4 <i>ānā</i>	৪	8 <i>ānā</i>	৮	12 <i>ānā</i>	১২	1 <i>rupayā</i>	১

- The *gaṇḍā* is written using digits and the BENGALI GANDA MARK precedes the unit (see Figure 1):

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

The *gaṇḍā* is divided into an intermediate unit of the *ānā* called the *pāī* (পাঁই), anglicized as “pie.” There are 5 *gaṇḍā* in 1 *pāī* and 4 *pāī* in 1 *ānā*. Thus, ৫ (5 *gaṇḍā*) = 1 *pāī*, ১০ (10 *gaṇḍā*) = 2 *pāī*, ১৫ (15 *gaṇḍā*) = 3 *pāī* and ২০ (20 *gaṇḍā*) = ১. As such, ১৩ “13 *gaṇḍā*” may also be expressed as “2 *pāī* and 3 *gaṇḍā*.”

Despite the common name, the Bengali *pāī* unit differs from the north Indic *pāī*. The Bengali *pāī* is closer to the north Indic *paisā* (পয়সা or পয়সী), anglicized as “pice.” The *paisā* is an intermediate unit of the *ānā*, which consists of 4 *paisā*.¹ The north Indic *pāī* is more similar to the Bengali *gaṇḍā*; there are 12 *pāī* in 1 *ānā* in the former system and 20 *gaṇḍā* in 1 *ānā* in the latter.²

- Historically, there is a unit smaller than the *gaṇḍā* called the *kaṛī* (কাড়ী), anglicized as “cowrie.” There are 4 *kaṛī* in 1 *gaṇḍā*. The *kaṛī* is written using BENGALI CURRENCY NUMERATOR FOUR and BENGALI CURRENCY NUMERATOR ONE LESS THAN THE DENOMINATOR. It is marked by the BENGALI CURRENCY DENOMINATOR SIXTEEN, which is written before the unit. Thus, 1 *kaṛī* is ০১ , 2 *kaṛī* is ০২ , 3 *kaṛī* is ০৩ , and 4 *kaṛī* is ০৪ .

When writing currency, only one mark is used. For isolated units, the relevant unit mark is written. For mixed units, the mark used is dependent upon the units to be expressed. If the *ānā* is present in a value that also has *rupayā* and/or *gaṇḍā*, then only the *ānā* mark is used: “15 *rupayā* and 3 *ānā*” is written as ১৫ , not as ১৫ or as ১৫ ; “27 *rupayā*, 6 *ānā*, and 5 *gaṇḍā*” is written as ২৭ . If a value contains only *ānā* and *gaṇḍā*, the BENGALI GANDA MARK is used: “5 *ānā* and 3 *gaṇḍā*” is written as ৫ (it may also be written without the *gaṇḍā* mark as ৫). If a value contains the *kaṛī* unit, the convention for writing that unit is followed: “5 *ānā* and 3 *kaṛī*” is written as ৫ ; “3 *gaṇḍā* and 2 *kaṛī*” is written as ৩ .

¹Figure 2 shows the use of the term “pice” to refer to the Bengali *pāī*.

²Halhed, 1778: 176–177.

Relationship to Other Systems The Bengali numeric notation system is closest to the north Indic system.³ Both are base-16 systems and both use the additive principle for writing fraction values, but they differ in the orthography and representation of fractions and unit marks. The Bengali system uses numerators for writing fraction values, while the north Indic system uses hash-like signs. Also, while currency units are written with distinctive signs in Bengali, the U+A838 NORTH INDIC RUPEE MARK is used for writing all units in the north Indic system. Thus, “3 *rupayā*” is ৩ in Bengali and ३ in Devanagari; “15 *ānā*” is ১৫ in Bengali and १५ in Devanagari. The north Indic *pāī* is written using fractions: ५/१ “5 *pāī*.” The equivalent unit in Bengal, the *gaṇḍā*, is written using digits and the BENGALI GANDA MARK: ৫ “5 *gaṇḍā*” (which, when converted to the north Indic system is equal to ५ “4 *pāī*”).

Modern Notation The use of currency marks and numerators in the Bengali script diminished in the latter half of the 20th century, when India changed its currency base. 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), or more commonly, *paisā*. In the new system there are 100 *paisā* in 1 *rupayā*, instead of the previous 16 *ānā* and 80 *gaṇḍā*. Modern currency is written using digits. The rupee mark was replaced with the Latin ‘Rs.’ U+20A8 RUPEE SIGN (‘Re.’ is used for a single rupee). The new rupee mark is written in Indic scripts with the syllable *ru* (Bengali রু), which is an abbreviation of *rupayā*. The U+09F3 BENGALI RUPEE SIGN is also used. When used, the modern Latin and Bengali rupee signs are written before currency values.

4 Proposal History

This proposal (ISO/IEC JTC1/SC2/WG2 N3311R L2/07-192R) 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” (ISO/IEC JTC1/SC2/WG2 N3311 L2/07-192). The UTC accepted L2/07-192 on August 9, 2007. The character was allocated at U+09FB, as proposed.⁴

5 References

- Beri, D. C. 19-?. हिन्दी बंगला टीचर [*Hindī Baṅgalā Ṭīcara* = Hindi Bengali Teacher]. Calcutta: Hindi Pracharak Pustakalaya.
- Chatterji, Suniti Kumar. 1927. *Bengali Self-Taught: By the Natural Method with Phonetic Pronunciation*. London: E. Marlborough & Co., Ltd.
- Everson, Michael. 2007. “Result of Repertoire Review for PDAMx of ISO/IEC 10646:2003 and future amendments.” ISO/IEC JTC1/SC2/WG2/N3321 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 Brasse. 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.” ISO/IEC JTC1/SC2/WG2 N3367 L2/07-354. October 7, 2007. <http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3367.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>.

³A description of the north Indic notation system is given in Pandey (2007).

⁴Everson, 2007; United States National Body, ISO, 2007: 2.

रुपया आना पाई लिखने का तरीका ।

২৥	৫	৯৥	১০	১২৥	১৫	১৯৥	১০	১২৥
୨)।)।=)।।)।।=)।।।)।।।=	୨	=)।
১৫	১১০	১১৫	১০	১০	১০	১২৥	১৫	১১০
୨)	୨)।	୨)।।	=)	≡)	।)	।)=	।)।	।)।।
১১৫	১০	১০	১০	১০	১০	১০	১০	১০
।)।।।	।)	।=)	।≡)	।।)	।।)	।।=)	।।≡)	।।।)
১০	১০	১০	১	১২৥	১৫	১১০	১১৫	১০
।।।)	।।।=)	।।।≡)	১)	১)=)	১)।	১)।।	১)।।।	১)
১০	১০	১০	১।।	১১০	১১০	১	১	১০০
১=)	১≡)	১।)	১।।)	১।।।)	১।।।≡)	১)	১০)	১০০)

Figure 1: The conversion and transliteration of Bengali currency notation to the north Indic system (from Beri, 19-?: 21). The specimen shows the ζ BENGALI GANDA MARK in print. The *ganḍā* unit is converted to the north Indic *pāṭ*. The specimen has two typographical errors. First, Bengali ১৫ is incorrectly transliterated as ২) (rows 3 and 4, column 1); the correct form is ২)। , as in the transliteration of ১১০ and ১১৫ . Second, Bengali ১২৥ is incorrectly transliterated as ১)=) (rows 7 and 8, column 5); the correct form is ১)। , without the second U+A838 NORTH INDIC RUPEE MARK, as in the transliteration of ১৫ , ১১০ , etc.

(b) Bengali Way of Writing Money.

One pice = ৫ i.e. 5 ‘ganḍā’ or 5 fours ($5 \times 4 = 20$) of cowrie-shells forming small currency in former times.

two pice = ১০ (10 ‘ganḍā’, or 40 cowrie shells).

three pice = ১৫ (15 ‘ganḍā’, or 60 cowrie shells).

one anna = ১০ (called ‘paṇ(a)’ or eighty cowries); two annas =

২০ (‘du-paṇ(a)’); three annas = ৩০ (tin(a)-paṇ(a)); three

annas one pice = ৩৫ ; three annas three pice = ৩১৫ ;

four annas = ১০ (called ‘cok(a), cauk(a)’ = group of four);

five annas = ১০ ; seven annas = ১০ ; eight annas = ১০

(‘du-cok(a)’ = two groups of four); ten annas = ১০ ; twelve

annas = ১০ (tin(a)-cok(a)); one rupee = ১০ ; Rs. 152, 13

annas, 3 pice = ১৫২১৫ ; Rs. 5, 3 pice = ৫১৫ , etc.

In Bengali books, the figures ১০ , ২০ , ৩০ , ১০ , etc., are used like the Roman numerals (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), etc., in European books: only the basis of computation is 8, not 10.

Figure 2: The orthography of currency in Bengali (from Chatterji, 1927: 196). The specimen shows the ζ BENGALI GANDA MARK in print.

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

1 ānā or $\frac{1}{16}$	/°	5 ānās	∕°	9 ānās	∥°	13 ānās	∩°
2 ānās	∕°	6 ānās	∣∕°	10 ānās	∥∕°	14 ānās	∩∕°
3 ānās	∨°	7 ānās	∣∨°	11 ānās	∥∨°	15 ānās	∩∨°
4 ānās or $\frac{1}{4}$	∣°	8 ānās or $\frac{1}{2}$	∥°	12 ānās or $\frac{3}{4}$	∩°		

Figure 3: 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 :—

/°—1 anna	∕°—5 annas	∥°—9 annas	∩°—13 annas
∕°—2 annas	∣∕°—6 annas	∥∕°—10 annas	∩∕°—14 annas
∨°—3 annas	∣∨°—7 annas	∥∨°—11 annas	∩∨°—15 annas
∣°—4 annas	∥°—8 annas	∩°—12 annas	∩∩°—1 rupee

Figure 4: 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.