Title: Proposal to Encode the Sign ANJI for Bengali
Source: Anshuman Pandey (pandey@umich.edu)
Status: Individual Contribution
Action: For consideration by UTC and WG2
Date: 2012-04-23

1 Introduction

This is a proposal to encode a new character in the Bengali block of the Universal Character Set (ISO/IEC 10646). This document is a minor revision of L2/11-359 and it replaces the earlier proposal. Properties of the proposed character are:

<table>
<thead>
<tr>
<th>Glyph</th>
<th>Code</th>
<th>Character Name</th>
<th>GC</th>
<th>CCC</th>
<th>BIDI</th>
<th>Mirrored</th>
<th>Linebreak</th>
</tr>
</thead>
<tbody>
<tr>
<td>וא</td>
<td>0980</td>
<td>BENGALI ANJI</td>
<td>LO</td>
<td>0</td>
<td>L</td>
<td>N</td>
<td>AL</td>
</tr>
</tbody>
</table>

The BENGALI ANJI is proposed for encoding at U+0980. This is the unassigned code-point at the beginning of the Bengali block, which is an appropriate allocation because the character is traditionally written at the beginning of texts as an invocation and at the head of the script.

2 Description

The Bengali sign АО (আঙী, আঙী < Skt. aŋji ‘ornament’) is written as an invocation at the beginning of texts. It is also called ঝুম্ম siddham. The character represents the Sanskrit phrase ঝুম্ম siddhirastu “may there be success”. In some interpretations, it is considered a symbol of the Hindu deity Ganesha in his aspect as patron of knowledge, represented specifically with his aṅkuśa “goad”, which symbolizes the notion of control or direction. For instance, in his Bengali Grammar, W. Yates, somewhat incorrectly, writes that the АО is a “contraction for the word Ganesh” (see figure 1). The sign is attested in inscriptions, manuscripts, and printed records.

D. C. Sircar offers the following details in Indian Epigraphy on the origins of the ANJI sign:

In inscriptions of the Gupta age, the auspicious word siddham at the beginning of epigraphs was sometimes represented by a crescent-like curved stroke opening towards the left. This mark resembles the symbol or mark of interpunctuation at the end of many of Aśoka’s edicts at Kalsī. It has been pointed out that some scholars regard it as standing for the pranava, i.e. the auspicious syllable om [...] there is some evidence to show that the symbol was read as siddham or siddhir=astu [...] The symbol was gradually modified to the shape of a stroke slanting towards lower right with its top curved or looped inwards to the right or slanting towards lower left with its top curved or looped inwards to the left. Sometimes the lower end of the stroke exhibits an ornamental downward flourish or upward bend. (1965: 94–95)

He also describes the emergence of the form used in Bengali and related scripts:
In the medieval inscriptions of Eastern India, the symbol resembles the Bengali numerical figure 7 (a vertical stroke with its top curved towards the left and ending in an inward bend) and it was used in the central parts of Bengal with the addition of a sign called *candra-bindu* or *ānunāsika* above. But, in the eastern-most areas of Bengal, the same symbol was written with the lower end curved considerably towards the right and was called *āñjī*, and the Bengali alphabet with this symbol at the head was called *āñjī-ka-kha*. (1965: 96)

Amitabha Bhattacharyya describes the *ANJI* as “a symbol which looks like either a *khaṇḍa* ta or an elephant’s trunk hanging down” (1994: 107).

Sircar’s description of the regional styles of writing *ANJI* are attested in the scripts related to Bengali. The Bengali *ANJI* corresponds to characters such as \(^{96} U+11480\) *TIRHUTA ANJI*, \(^{97} U+AAF2\) *MEITEI MAYEK ANJI*, and the Newar sign *siddhi*. There may be a corresponding character in Devanagari that is used in the transliteration of Bengali (see figure 2); however, additional research is required in order to ascertain its usage.

3 Character Data

*Variants* The *ANJI* was written differently throughout the Bengal region. In some cases a *candrabindu* was placed above the sign: ₃. This form ₃ may be represented using the sequence `<♀ ANJI, ♀ BENGALI SIGN CANDRABINDU>`. Other variants are to be managed through fonts.

*General Category* The character is assigned ‘gc=Lo’ because it may be combined with marks such as *CANDRABINDU*.

*Collation* The *ANJI* should be sorted at the head of the script before ☞ U+0985 BENGALI LETTER A.

*‘Confusable’ Data* The ☞ ANJI resembles ☞ U+09ED BENGALI DIGIT SEVEN.

*Annotation to the Names List* The following annotation is to be added to the names list:

```
@  Sign
0980  BENGALI ANJI
     = siddham, siddhirastu
     * used at the beginning of texts as an invocation
```

4 References


Some consonants, when combined with this form of ः, assume a peculiar shape; as,
क्रा, क्र, त्रा, त्र, न्द्रा, न्द्र.
The vowel ः combined with ः, is thus expressed: क्रु।
The vowel ः combined with ः, is thus expressed: क्रु।
So क्रू, क्रृ, श्रु, श्रृ, द्रू, द्रृ।
Sometimes the vowel ः has this form ( ) ; as,
ँ गु, शु, न्तु, श्तु।
Remark also ः हु, and ः ह्रि, and ः (for ः) त।
In addition to the above, which are very common, a few other compound letters are worthy of notice, as they are frequently used by the natives, although generally avoided in good printing.
ः for ः tu।ः for ः ku।
ः for ः mu।ः for ः hra।
(*) over a letter signifies that it is nasal.

The following is a contraction for the word Ganesh.

is a contraction for the name of God.

Section VII.—Of the Permutation of the Letters,
(Sandhi.)*

The changing of the letters इ and ई to ए; of उ and ऊ to ऋ; of ऋ and ऌ to ओ, is called gun (pron. goon.)
The changing of the letters ओ to आ; इ ई and ए to ऐ; उ ऊ and ऋ to न; and ऋ ओ to ओ, is called briddhi.

* This section may be entirely omitted by the learner on going through the grammar the first time, and studied afterwards, in connexion with the chapter on compound words. As it strictly belongs to orthography, order requires its insertion in this place.
TEXT IN NĀGARĪ SCRIPT

(1) q सिद्धिरस्तु | विक्रमाधिय-न्यूथे: | संवत् 1297 | येरूँट-बरि १५ | ज्ये भारसेत्र-ध्वाल | अनां-पिन्तावल्ले | संवत् १२९७ | निर्दर्षित

(2) वेरूँट-बरि | द्वारा-प्रजाश्रयो | राजा | कुस्मेश्वरराज | अनां-पिन्तावल्ले | संवत् १२९७ | निर्दर्षित

(3) येरूँट-बरि | द्वारा-प्रजाश्रयो | राजा | कुस्मेश्वरराज | अनां-पिन्तावल्ले | संवत् १२९७ | निर्दर्षित

TRANSLATION

[Symbol meaning ‘Let there be success.’] Let there be success. On Thursday, the 15th lunar day of the dark half of Jyeṣṭha in the year 1297 of king Vikramāditya, when Suratāṇa Mauijadin, who is the forest-fire to the grass and dry wood that are the numerous powerful enemies, was ruling the kingdom, the illustrious Mantreśvara-Kāmadeva who is the son of Ajayapāla, grandson of Uttara and great-grandson of Jagatpāla and belongs to the Ajihila‘Ajihina)-gotra and to the Sopatha family that is the foremost clan of the Kṣatriya community, performed Gayā (i.e. the śrāddha of his ancestors at Gayā together with his wife. Likewise, his brother Rājadeva, father of Jaitrapāla, also performed Gayā (i.e. the śrāddha probably of Jaitrapāla, at Gayā). The witness [to the performance of the ceremonies] is Prapitāmaha (i.e. the god Prapitāmaheśvara of Gayā). Let there be welfare for the world. [The people whose śrāddha has been performed] at Gayā have obtained their deliverance [from their sins or the effects of their karman].

ANSARI’S TRANSCRIPT AND TRANSLATION OF THE RECORD

TEXT

1 Om siddhi rastu / vikramadityayanipateh / samvat 1257 jyeṣṭha vadi 15 ravau / Pravalatarani Kṛ-
2 ta ṭpa dvijašvakṣa dāvānalo Śrī Suratāṇa Mojitina rājyaṁ kurveṇe Nṛ-
3 pa pravara kulsoya dhanārththayā Nājahila gotriṇe Śrī Jagatpāla Prapitāmāhena /
4 ...? rapitāmāhena / Ajayapāla Pitrakena / Śrī Mantreśvara Kāmadeva Gaja-
5 tā sapatnikena / tadā (yā?) trāitre (tathā bhrātrā?) Bājadeva Jaitrapāla labdhena Gayākṛta
6 Prapitāmahakṣāyāhni / sā janjava (ṣrūtabhā?) śrī (vato?) jagatagho (jagatottha?) Gayāsu hatā (Gayā uddhāta?) / matā (gatā?) /

Figure 2: Excerpt showing the use of U+0071 LATIN SMALL LETTER Q for representing 'q' ANJI in Devanagari text (from Sircar 1979: 39). The sign appears at top in the Devanagari as “q” in the phrase “q सिद्धिरस्तु” and is glossed as “[Symbol meaning ‘Let there be success.’]”.
### A. Administrative

1. **Title:** Proposal to Encode the Sign ANJI for Bengali

2. **Requester's name:** Anshuman Pandey (pandey@umich.edu)

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

4. **Submission date:** 2012-04-23

5. **Requester's reference (if applicable):**

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

### B. Technical – General

1. **Choose one of the following:**
   - a. This proposal is for a new script (set of characters):
   - Proposed name of script:
   - b. 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

3. **Proposed category (select one from below - see section 2.2 of P&P document):**
   - A-Contemporary
   - B.1-Specialized (small collection)
   - B.2-Specialized (large collection)
   - C-Major extinct
   - D-Attested extinct
   - E-Minor extinct
   - F-Archaic Hieroglyphic or Ideographic
   - G-Obscure or questionable usage symbols

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. **Fonts related:**
   - a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
     - Anshuman Pandey
   - b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): Anshuman Pandey (pandey@umich.edu)

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:**
   - 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. **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](http://www.unicode.org) for such information on other scripts. Also see Unicode Character Database ([http://www.unicode.org/reports/tr44/](http://www.unicode.org/reports/tr44/)) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

---

## C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?
   - **No**
   - If YES explain

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**
   - If YES, with whom?
   - If YES, available relevant documents:

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?
   - **Yes**
   - Reference: Size of user community is unknown.

4. The context of use for the proposed characters (type of use; common or rare)
   - **Common**
   - Reference: Character used as invocation symbol in inscriptions, manuscripts, and other records

5. Are the proposed characters in current use by the user community?
   - **Yes**
   - Reference: Character is historical, but may be in current use

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?
   - **Yes**
   - If YES, is a rationale provided?
     - Reference: Character proposed for encoding in the Bengali block of BMP

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

8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?
   - **No**
   - If YES, a rationale for its inclusion provided?

9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?
   - **No**
   - If YES, a rationale for its inclusion provided?

10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?
    - **Yes**
    - If YES, is a rationale for its inclusion provided?
     - Reference: Corresponds to characters used for similar purposes in other scripts

11. Does the proposal include use of combining characters and/or use of composite sequences?
    - **No**
    - If YES, is a rationale for such use provided?

12. Does the proposal contain characters with any special properties such as control function or similar semantics?
    - **No**
    - If YES, describe in detail (include attachment if necessary)

13. Does the proposal contain any Ideographic compatibility characters?
    - **No**
    - If YES, are the equivalent corresponding unified ideographic characters identified?