

**\*\*\*\*This proposal has not been submitted\*\*\*\***  
**\*\*\*This document is displayed for initial feedback only\*\*\***  
**\*\*\*This proposal is currently incomplete\*\*\***

L2/03-089

**ISO**  
**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**  
**ORGANISATION INTERNATIONALE DE NORMALISATION**

ISO/IEC JTC 1/SC 2/WG 2

Universal Multiple-Octet Coded Character Set  
(U C S)

ISO/IEC JTC1/SC2/WG2 N????  
Date: 200?-??-??

**Title:** Proposal for addition of BENGALI LETTER OPEN A & AND BENGALI LETTER CENTRAL E

**Source:** Andy White

**Status:** To be submitted ASAP

**Action:** To be corrected, updated, completed and submitted ASAP

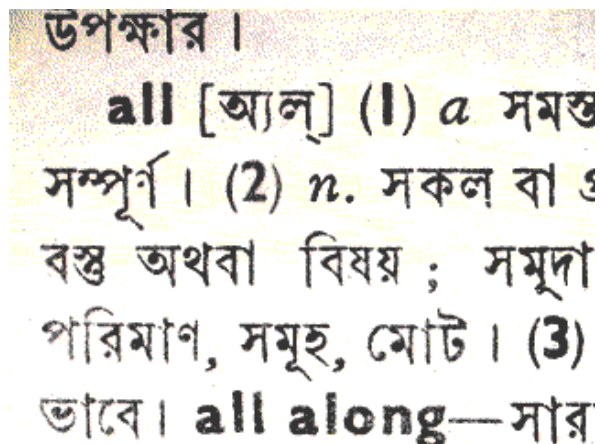
### Summary

This is a proposal for 'অ্য' BENGALI LETTER OPEN O & 'এ্য' BENGALI LETTER CENTRAL E. (*Names of these letters are currently subject to change.*) These letters are being proposed as part of a series of proposals aimed at bringing the Bengali code block in line with the Devanagari block.

### Introduction

The Devanagari code block includes the letters CANDRA E (U+090D) & CANDRA O (U+0911). These letters are used to transcribe the sounds 'e' as in 'err' & 'a' as in 'ball'. Bengali has equivalent letters to these, and so they can be placed at the equivalent positions in the Bengali block. In fact the 'equivalent positions' were left free especially for this event.

Examples of these characters in use are to be given below.



উপক্ষার ।  
all [অ্যন্] (1) a সমস্ত  
সম্পূর্ণ। (2) n. সকল বা ও  
বস্তু অথবা বিষয় ; সমুদা  
পরিমাণ, সমূহ, মোট। (3)  
ভাবে। all along—সার

**The above example shows the 'a' in the word 'all' as being transcribed with the proposed BENGALI LETTER OPEN O**

**Image taken from page 21 of 'Samsad English- Bengali dictionary' - S. Biswas 1999.**

*If anyone has a copy of A.T. Dev's English to Bengali Dictionary, please could you scan in an example of the EYA vowel for me - thanks.*

It has been proposed that the above letters could be encoded as 'Vowel+Virama+Ya...' sequences. This proposal rejects this idea as:

1. The sequence 'Vowel+Virama+Ya...' is illogical to scholars of Bengali and indeed Indic languages in general.
2. Such sequences are not semantically equivalent to the intended
3. There are no other cases of a Vowel+Virama combination in the Unicode encoding model.
4. Yaphalaa is not equivalent to 'Virama+Ya'
5. ISCII implementations encode these letters as separate characters corresponding to the Devanagari Candra A & E. Unicode should follow the example of these implementations.
6. Other good reasons to be added soon!

## **Collation**

Collation of the proposed will be as per the Devanagri equivalents (Further explanation to be added)

## **Implications on current rendering systems.**

Current rendering systems should have no problems displaying the proposed letters as the proposed do not take part in conjunct formations etc. (i.e. they need no further complex script processing)

## **Interoperability with other standards**

The proposed letters are undefined in ISCII. However, ISCII implementations encode there letters at the positions allocated to Devanagari Candra E & Candra O (as in this proposal)

## **PROPOSAL SUMMARY**

*ISO/IEC JTC 1/SC 2/WG 2*

### **PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646**

*Please fill all the sections A, B and C below.*

(Please read Principles and Procedures Document for guidelines and details before filling this form.)

See <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html> for latest Form.

See <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for latest Principles and Procedures document.

See <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest roadmaps.

(Form number: N2352-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09)

## **A. Administrative**

1. Title: **Proposal for addition of BENGALI LETTER OPEN A & AND BENGALI LETTER CENTRAL E**

2. Requester's name: **Andy White**

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

4. Submission date: **ASAP \*\*\***

5. Requester's reference (if applicable):

6. (Choose one of the following:) This is a complete proposal: **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:)

a. This proposal is for a new script (set of characters): **No**

Proposed name of script:

b. The proposal is for addition of character(s) to an existing block: **Yes**

Name of the existing block: **BMP**

2. Number of characters in proposal: **Two**

3. Proposed category (see section II, Character Categories): **\*\*\***

4. Proposed Level of Implementation (1, 2 or 3)

(see clause 14, ISO/IEC 10646-1: 2000): **Any level is acceptable**

Is a rationale provided for the choice? \_\_\_\_\_

If Yes, reference: \_\_\_\_\_

5. 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 ISO/IEC 10646-1: 2000? **Yes**

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

6. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? **The proposer**

If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used: **To be identified**

7. 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**

8. 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**

9. 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 Unicode Character Database

<http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html>

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.)? **Yes, user groups**

If YES, with whom? **Bengali Font project group, Bengali Linux Group etc., list to be supplied\*\*\***

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? **No**

Reference: **Indic Community (among others)**

4. The context of use for the proposed characters (type of use; common or rare) **Common\_**

Reference: **As enclosed**

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

If YES, where? Reference: **\*\*\***

6. After giving due considerations to the principles in *Principles and Procedures document* (a WG 2 standing document) must the proposed characters be entirely in the BMP? **Yes**

If YES, is a rationale provided? **\*\*\*To be provided\*\*\***

If YES, reference: **enclosed**

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

8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? **No but see proposal body\*\*\***

If YES, is a rationale for its inclusion provided?

If YES, reference: **\_**

9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? **possibly, see poposal body\*\*\***

If YES, is a rationale for its inclusion provided?

If YES, reference:

10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? **No**

If YES, is a rationale for its inclusion provided?

If YES, reference:

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)? **\*\*\*Not Sure\*\*\***

If YES, is a rationale for such use provided?

If YES, reference:

Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? **\*\*\*\*\***

If YES, reference:

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 character(s)? **No**

If YES, is the equivalent corresponding unified ideographic character(s) identified?

If YES, reference:

## A.1 Submitter's Responsibilities

The national body or liaison organization (or any other organization or an individual) proposing new character(s) or a new script shall provide:

1. Proposed category for the script or character(s), character name(s), and description of usage.
2. Justification for the category and name(s).
3. A representative glyph(s) image on paper:  
If the proposed glyph image is similar to a glyph image of a previously encoded ISO/IEC 10646 character, then additional justification for encoding the new character shall be provided.  
**Note:** Any proposal that suggests that one or more of such variant forms is actually a distinct character requiring separate encoding, should provide detailed, printed evidence that there is actual, contrastive use of the variant form(s). It is insufficient for a proposal to claim a requirement to encode as characters in the Standard, glyphic forms which happen to occur in another character encoding that did not follow the Character-Glyph Model that guides the choice of appropriate characters for encoding in ISO/IEC 10646.  
**Note:** WG 2 has resolved in Resolution M38.12 not to add any more Arabic presentation forms to the standard and suggests users to employ appropriate input methods, rendering and font technologies to meet the user requirements.
4. Mappings to accepted sources, for example, other standards, dictionaries, accessible published materials.
5. Computerized/camera-ready font:  
Prior to the preparation of the final text of the next amendment or version of the standard a suitable computerized font (camera-ready font) will be needed. Camera-ready copy is mandatory for final text of any pDAMs before the next revision. Ordered preference of the fonts is True Type or PostScript format. The minimum design resolution for the font is 96 by 96 dots matrix, for presentation at or near 22 points in print size.
6. List of all the parties consulted.
7. Equivalent glyph images:  
If the submission intends using composite sequences of proposed or existing combining and non-combining characters, a list consisting of each composite sequence and its corresponding glyph image shall be provided to better understand the intended use.
8. Compatibility equivalents:  
If the submission includes compatibility ideographic characters, identify the equivalent unified CJK Ideograph character(s).
9. Any additional information that will assist in correct understanding of the different characteristics and linguistic processing of the proposed character(s) or script.

○○○›☺