

TITLE: Proposal to add three Kaithi punctuation marks
SOURCE: Deborah Anderson, Script Encoding Initiative, UC Berkeley
STATUS: Liaison contribution
ACTION: For consideration by WG2
DISTRIBUTION: ISO/IEC JTC1/SC2/WG2
DATE: 15 October 2008

After consultation with the Kaithi proposal author, Anshuman Pandey, the following characters are proposed for inclusion in Amendment 6.2:

110BD KAITHI NUMBER SIGN



110C0 KAITHI DANDA

110C1 KAITHI DOUBLE DANDA



The glyphs also appear in document N3546.

The background document, N3389, is the full Kaithi proposal and it includes published examples of these characters. (Note: Document N3520 is an interim report announcing that potential new characters will be proposed. Document N3520 does not include the dandas, but after discussing the new danda policy in the Principles and Procedures document N3452, A. Pandey agreed they should be separately encoded.)

The KAITHI NUMBER SIGN is being separately encoded from KAITHI ENUMERATION SIGN, an inline spacing character, because rendering engines will need to handle them differently. KAITHI NUMBER SIGN should work in a way that is similar to 0600 ARABIC NUMBER SIGN.

Character properties:

110BD;KAITHI NUMBER SIGN;Cf;0;AN;;;;;N;;;;;

110C0;KAITHI DANDA;Po;0;L;;;;;N;;;;;

110C1;KAITHI DOUBLE DANDA;Po;0;L;;;;;N;;;;;

**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 three additional Kaithi punctuation characters**

2. Requester's name: *Deborah Anderson, Script Encoding Initiative, UC Berkeley*

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

4. Submission date: *15 October 2008*

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: *Kaithi*

2. Number of characters in proposal: *3*

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

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

6. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *see N3389*

b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? *see N3389*

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,
please
also see
N3389*

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.

¹ Form number: N3152-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, 2008-05)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain	<i>no</i> <i>(please see discussion in N3389 and this document)</i>
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.)? If YES, with whom? If YES, available relevant documents:	<i>no</i>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	<i>see N3389</i> <i>Awadhi, Bhojpuri, Magahi, and Maithili speakers; as well as linguists, historians, legal scholars working with sources from colonial South Asia.</i>
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>common</i> <i>Court records from colonial India, pedagogical materials from north India, commercial and accounting records; religious and literary texts; bibles printed in north India during the 19th and early 20th century.</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>see N3389</i> <i>In India, the United States, and other localities.</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:	<i>no</i>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	<i>n/a</i>
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>no</i>
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>no</i>
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>yes</i> <i>yes</i> <i>see N3389</i>
11. Does the proposal include use of combining characters and/or use of composite sequences? 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:	<i>no</i>
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)	<i>yes</i> <i>see N3389</i>
13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? If YES, reference:	<i>no</i>