


Title: Proposal to Encode the SUTRA MARK for Sharada
Source: Script Encoding Initiative (SEI)
Author: Anshuman Pandey (pandey@umich.edu)
Status: Liaison Contribution
Action: For consideration by UTC and WG2
Date: 2012-05-03


1 Introduction

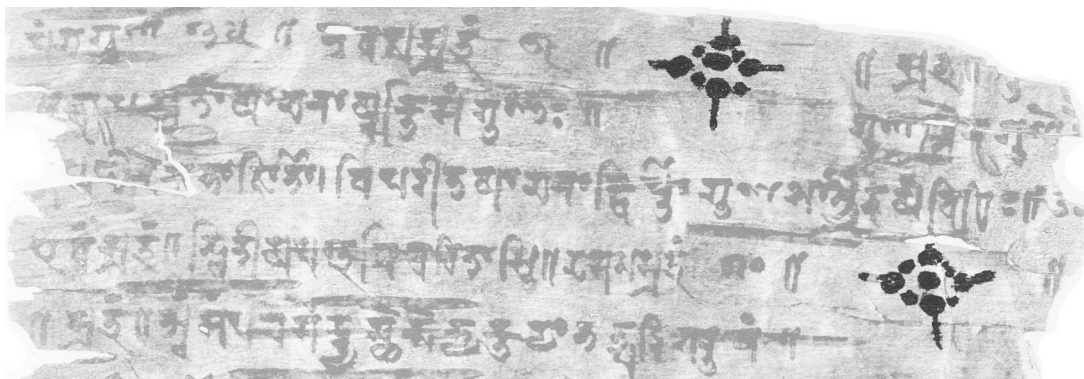
This is a proposal to encode a new character in the Sharada block of the Universal Character Set (ISO/IEC 10646). Properties of the proposed character are:

GLYPH	CODE	CHARACTER NAME	GC	CCC	BIDI	MIRRORED	LINEBREAK
	111CD	SHARADA SUTRA MARK	Po	0	L	N	BB

This character was identified in the original proposal to encode Sharada in the UCS (N3595 L2/09-074R2: 9–10), but it was not included in the proposed character repertoire because additional research regarding its usage was required.

2 Description

The  SHARADA SUTRA MARK is used for indicating the end of a *sūtra*, or ‘rule’:



The proposed character is used throughout the *Bakhshali Manuscript*, a mathematical treatise written in Sharada, which is dated to the 2–4th centuries CE.

In his edition of the *Bhakshali Manuscript* (1927), G. R. Kaye shows two different glyphs for the proposed character (see figures 3 and 4). The difference is in the number of dots. As shown above, the SUTRA MARK has five large and four small dots. The glyph shown in figure 3 lacks the smaller dots and does not occur in the manuscript or in the transliteration; it is likely a typographic error. The nine-dot form is the correct form.

3 Character Data

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

```
@      Punctuation
111CD  SHARADA SUTRA MARK
      = used at the end of a sutra in the Bakhshali manuscript
```

4 References

Kaye, G. R. 1927. *The Bakhshālī Manuscript: A Study in Mediæval Mathematics*. Archaeological Survey of India, New Imperial Series, Vol. XLII. Parts I & II, Part III. Calcutta: Government of India Central Publication Branch.

Pandey, Anshuman. 2009. "Proposal to Encode the Sharada Script in ISO/IEC 10646". N3595 L2/09-074R2. August 5, 2009. <http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3595.pdf>

5 Acknowledgments

This project was made possible in part by a grant from the United States National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at the University of California, Berkeley). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment for the Humanities.

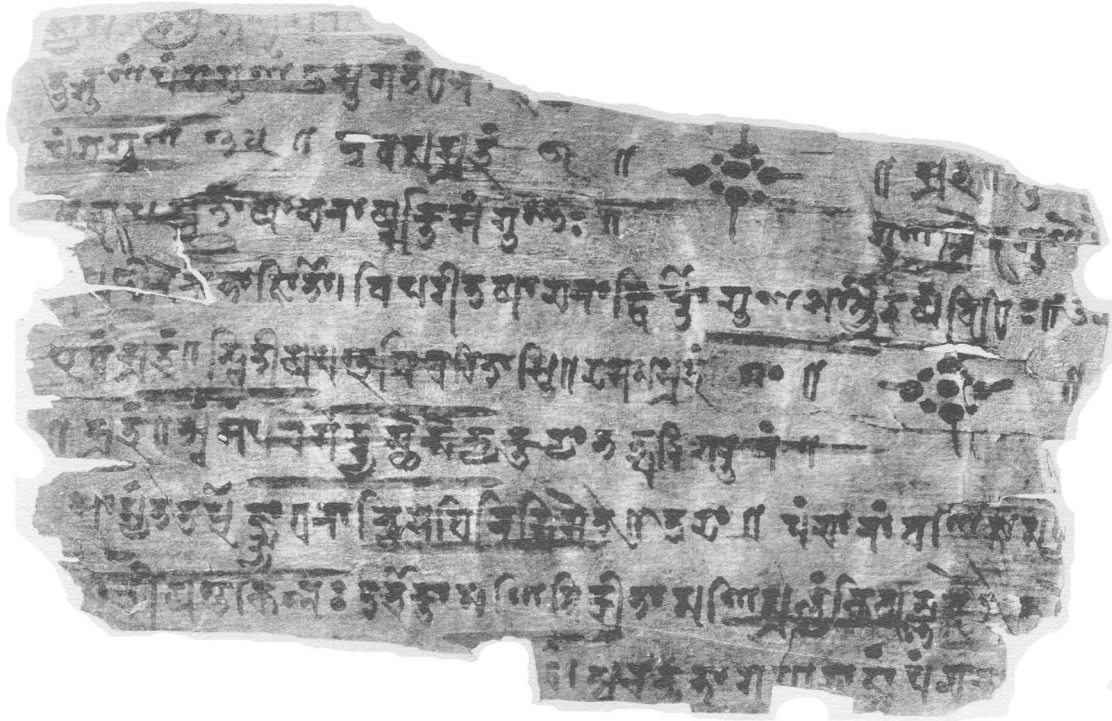


Figure 1: Use of the SUTRA MARK in folio 1r of the manuscript (from Kaye 1927: 22).

bhyāsa i rū [1r.]
 turguṇaṃpañchaguṇaṃhastagatamdhanam | ja
 pañchaguṇaṃ 25 || navamasūtraṃ 9 || ⋮
—•—
⋮ || sūtraṃ || guṇaupri
 thagrūpayutaūyāchanāyuktisaṃguṇauḥ || guṇanenagune . ai
 rūpahinenabhājitau | viparītayāchanākshiptauguṇasāsyorayamvidhiḥ || . .
 evaṃsūtraṃ || dvitīyapatrevivaritāsti || daśamasūtraṃ 10 || ⋮
—•—
⋮ ||
 || sūtraṃ || aṃśāṃviśodhychchedebhyakuryātatparivartanaṃ || . .
 sāsyamtataproyjyadhanānviśavinirdiśet || udā | pañchānāmvaṇijāmadhye
vikriyatekilah tatroktāmaṇivikritāmaṇimūlyamkiyadbhaves
 dam | ardhatribhāgapādāṅśampañchabhā

Figure 2: Use of the SUTRA MARK in a transliteration of folio 1r (from Kaye 1927: 105).

54. The text consists of rules (*sūtras*) and examples. There is no explanation whatever of the processes by which the rules were obtained, that is, there is no mathematical theory at all. In this the work follows the usual Indian fashion, as exhibited in all early texts. But there is a good deal of mathematical theory implied, and the rules and examples are often set forth in such a way as to convey the principles followed quite clearly to the student.

55. The rules (*sūtras*) are written in verse and are generally numbered; and often in the solutions of the examples phrases from the *sūtras* are quoted. When Dr. Hoernle tried to re-arrange the leaves of the manuscript, he took the numbered *sūtras* as the basis of his order, but they were too few in number to lead to a satisfactory result. These *sūtras* do not represent, as might be expected, the most valuable part of the text. They are usually of particular application rather than general and are often very obscurely expressed.

56. The examples given are generally formally stated in full—without the use of notation or abbreviation of any kind; and in most cases they are stated in verse. They are introduced by the term *udā*, an abbreviation for *udāharanam* “an example.” After the question sometimes comes a formal statement with numerical symbols and abbreviations, often arranged in cells. Then comes the solution or working (*karanā*), and here, sometimes, fragments of the *sūtras* are quoted. Finally come demonstrations—often more of the nature of verification than proof. Generally these demonstrations, by the aid of the answer found to the question, rediscover one of the original elements of the problem; and sometimes several such demonstrations are attached to an individual problem, but sometimes the variation is merely a matter of the form of statement.

The full scheme of exposition is therefore—


Sūtram or rule.
Udāharanam or “example”: indicated by *udā*°.
Sthāpanam or “statement.”
Karanam or “solution.”
Pratyayam or “verification.”

The end of each *sūtra* is marked after the last example by the device and the number of the *sūtra* is also given at the end.



Figure 3: Description of use of the SUTRA MARK (from Kaye 1927: 22). The graphical form of the character on this page differs from that shown in figure 4, which is the form used throughout the manuscript and the transliteration. The form on this page, which does not contain the four smaller dots, is likely an error.

Punctuation.

Besides the visarga the single bar | and double bar || are used, and the end of a *sūtra* is marked by . The *virāma* appears to be used to mark a pause or the end of a verse. On folio 5 recto two punctuation signs, which are possibly cancellation marks, are employed.

Cancellation.

On several occasions letters and words are cancelled by marks like accents placed above the letters (See folios 8 and 9).

Figure 4: Another description of the SUTRA MARK, showing the normative form of the glyph (from Kaye 1927: 99).

**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 SUTRA MARK for Sharada
2. Requester's name:	<i>Script Encoding Initiative (SEI) / Anshuman Pandey (pandey@umich.edu)</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Liaison contribution</i>
4. Submission date:	<i>2012-05-03</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<input checked="" type="checkbox"/> Yes
(or) More information will be provided later:	<input type="checkbox"/>

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:	<i>Sharada</i>
2. Number of characters in proposal:	<i>1</i>
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary <input checked="" type="checkbox"/> B.1-Specialized (small collection) <input type="checkbox"/> B.2-Specialized (large collection) <input type="checkbox"/>	
C-Major extinct <input type="checkbox"/> D-Attested extinct <input type="checkbox"/> E-Minor extinct <input type="checkbox"/>	
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/> G-Obscure or questionable usage symbols <input type="checkbox"/>	
4. Is a repertoire including character names provided?	<input checked="" type="checkbox"/> Yes
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<input checked="" type="checkbox"/> Yes
b. Are the character shapes attached in a legible form suitable for review?	<input checked="" type="checkbox"/> Yes
5. Fonts related:	
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Anshuman Pandey</i>
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>Anshuman Pandey (pandey@umich.edu)</i>
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<input checked="" type="checkbox"/> Yes
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<input checked="" type="checkbox"/> 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)?	<input checked="" type="checkbox"/> 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> for such information on other scripts. Also see Unicode Character Database (<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.

¹ Form number: N3902-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, 2009-11, 2011-03)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain	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.)? If YES, with whom? If YES, available relevant documents:	No
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	Yes <i>Size of user community is unknown.</i>
4. The context of use for the proposed characters (type of use; common or rare) Reference:	Common <i>Character used in historical Sharada manuscripts</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	Yes <i>Character is historical, but is attested in manuscripts and printed works, which are used by modern scholars</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:	N/A
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? If YES, is a rationale for its inclusion provided? If YES, reference:	No
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:	No
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:	No
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:	No
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)	No
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:	No