

Title: Proposal to Encode the Sign SIDDHAM for Devanagari
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 Universal Character Set: ॐ DEVANAGARI SIDDHAM. Properties of the proposed character are:

GLYPH	CODE	CHARACTER NAME	GC	CCC	BIDI	MIRRORED	LINEBREAK
ॐ	A8FC	DEVANAGARI SIDDHAM	So	0	L	N	BB

The character is proposed for encoding in the ‘Devanagari Extended’ (U+A8E0) block at U+A8FC, which is the next available code point in the block.

2 Description

The Devanagari sign ॐ SIDDHAM (Sans. सिद्धम् *siddham* “accomplished”) is written as an invocation at the beginning of texts. It represents the Sanskrit phrase सिद्धिरस्तु *siddhirastu* “may there be success”. In use, the sign is often followed by the word स्वस्ति *svasti* “may it be auspicious”, ie. ॐ स्वस्ति, which is expressed as “*siddham svasti*”. The SIDDHAM sign is attested in inscriptions, manuscripts, and printed records.

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

The auspicious word *siddham* is found at the beginning of numerous early records. This word was in later times generally indicated by a symbol of varying shape. The symbol was sometimes followed by the auspicious word *svasti* which, in some cases, stands singly at the beginning of epigraphic records. [...] In some cases again a symbol, which is apparently a variant of the *siddham* sign, appears at the end of a document and, in rare cases, also in the body of its text especially at the end of a section of it. [...] There is no doubt that the symbol stood for the word *siddham* found at the beginning of numerous early inscriptions. [...] 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. (1965: 92–95)

3 Character Details

Representative Glyph Shape The printed form of the sign has several variants. Müller and Nanjio show the ‘closed’ glyph ॐ (see figure 1) and Bendall shows the ‘open’ ॐ (see figure 2). The latter is more common. For this reason the form ॐ is proposed as the representative glyph shape. It is a sort of graphical average between the ‘closed’ and ‘open’ forms.

Correspondences The sign corresponds to characters in other scripts, such as ཨ U+0FD3 TIBETAN MARK INITIAL BRDA RNYING YIG MGO MDUN MA and the sign ཨྱ *siddham* of the Siddham script.

Collation The sign SIDDHAM is sorted at the head of the script before अ U+0905 DEVANAGARI LETTER A.

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

@	Sign
A8FC	DEVANAGARI SIDDHAM
=	siddhirastu
*	used at the beginning of texts as an invocation

4 References

- Bendall, Cecil. 1886. *A Journey of Literary and Archæological Research in Nepal and Northern India, during the Winter of 1884-5*. Cambridge: Cambridge University Press.
- Müller, Friedrich Max and Bunyiu Nanjio [eds]. 1884. *The Ancient Palm-Leaves*. Containing the Pragñâ-pâramitâ-hridaya-sûtra and the Ushnîsha-vigaya-dhâranî. *Anecdota Oxoniensia*: Texts, documents, and extracts chiefly from manuscripts in the Bodleian and other Oxford libraries. Aryan Series, Vol. I, Pt. III. With an appendix by Georg Bühler. Oxford: Clarendon Press.
- Sircar, D. C. 1965. *Indian Epigraphy*. Delhi: Motilal Banarsidass.

नम सर्वज्ञय
 ॐ नम सर्वज्ञय ॥
 ॐ नमस्सर्वज्ञाय ॥

अर्यवरोक्तेश्वरा बोधिसत्त्व खंविरेय प्रज्ञापरमितय चर्ये
 आर्यावलोकितेश्वर बोधिसत्त्व गम्भीरय प्रज्ञपरमितयं चर्ये
 आर्यावलोकितेश्वर बोधिसत्त्वो गम्भीरं प्रज्ञपारमितायं चर्ये

चरमणे च्यम्भरोकयति स्म पंच स्कन्धस्थश स्वभावशुन्य
 चारमनुष्य व्यवलोकयति स्म पंच स्कन्धास्तश्च स्वभावशुन्यां
 चरमानो व्यावलोकयति स्म पंच स्कन्धास्ताश्च स्वभावशुन्यं

पश्यति स्म ईह शरिपुत्र रूपं शुचिता शून्यतैव रूपं न
 पश्यति स्म इह शरिपुत्र रूपं शुन्यता शून्यतेव रूपं रूपं न
 पश्यति स्म इह शारिपुत्र रूपं शुन्यता शून्यतैव रूपं रूपान्न

पृथङ्गशून्यत शून्यधाय त पृणश्रूपं पंदूप स शून्यत य शून्यता
 पृथक्शून्यत शून्यतय न पृथक्श्रूपं यदूपं स शून्यत य शून्यते
 पृथक् शून्यता शून्यताया न पृथग्यूपं यदूपं सा शून्यता य शून्यता

वदूपं एवपेव वेदनसंस्हरविज्ञतनि
 वदूपं एवंपथ वेदनसंज्ञसंस्कारविज्ञननि च
 तदूपं एवमैव वैदनसंज्ञसंस्कारविज्ञनाति

Figure 1: Use of SIDDHAM at the beginning of a Devanagari transliteration of a Siddham manuscript (from Müller and Nanjio 1884: 28).

“Samvat 151, on the second day of the light half of Vaiçākha, Bhojamati wife of Atitalambha gave two mās [of land?] to the Pañch-committee of Lañjagval, together with a water-receptacle, for their perpetual enjoyment thereof.”

What *jaladrona* may mean precisely I have no means of telling; and the dictionaries give no help. I at first thought from the position of the stone and from a possible connexion with *√dru* ‘run’ that it must mean water-course, like *pranāli*: but the ordinary meaning of *drona*, ‘tub’, rather suggests a reservoir; and to this view Dr Bühler, I find, is inclined. The word *drona* occurs also in the next inscription, and there the first meaning is perhaps more probable. The gradual approximation to Kuṭila forms in the characters of this inscription is noticeable, particularly in the lengthened and more sweeping *curves* of medial *ā* and *ī*.

V. Dedicatory verses on the pedestal of a figure of the sun-god, Patan, Nepal. Dated [Nepal] Samvat 203 (A.D. 1083).

See pp. 8–9 with plate. Space covered by inscription, $5\frac{1}{2} \times 2\frac{3}{4}$ inches.

1. ७ त्रिभिर्वर्षैः समायुक्ते संवत्सरशतदये । वैशाखशु
2. क्लृप्तम्यां [sic] बुधे पुष्योदये शुभा [॥] श्रीचणोदेवभूनायत
3. नयो धर्मतत्परः । श्रीवाणदेवः कृतवान् प्रतिमां सु
4. प्रतिष्ठित[॥] । दिवाकरस्य या मात्रा पुरा संकल्पिता मुदा
5. कर्तुंस्तेनास्तु सततं तेजोवृद्धिरनुत्तरा ॥ ० ॥

Translation.

When two hundred years were joined with three, on the 7th of the bright half of Vaiçākha, on Wednesday, Pushyā was auspicious at its rising. Vāṇadeva son of king Yaçodeva, religiously disposed, made [this] image well set up in honour of the Sun, which had previously been planned by his mother with

Figure 2: Use of SIDDHAM at the beginning of Devanagari transliteration of a Newar inscription (from Bendall 1886: 80).

**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 Sign SIDDHAM for Devanagari
2. Requester's name:	<i>Anshuman Pandey (pandey@umich.edu)</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual contribution</i>
4. Submission date:	<i>2012-04-23</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<i>Yes</i>
(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:	<i>Devanagari Extended</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?	<i>Yes</i>
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<i>Yes</i>
b. Are the character shapes attached in a legible form suitable for review?	<i>Yes</i>
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?	<i>Yes</i>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>Yes</i>
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)?	<i>Yes</i>

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 as invocation symbol in inscriptions, manuscripts, and other records</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	Yes <i>Character is historical, but may be in current use</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:	Yes Yes <i>Proposed for encoding in the Devanagari Extended block of BMP</i>
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:	Yes Yes <i>Corresponds to characters used for similar purposes in other scripts</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:	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