1 Introduction

This is a proposal to encode a new character in the Sharada block of the Universal Character Set (ISO/IEC 10646). Sharada will be included in a future version of the UCS at the range U+11180..U+111DF (see N3595 L2/09-074R2). 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>आठ</td>
<td>111C9</td>
<td>SHARADA EKAM</td>
<td>Lo</td>
<td>0</td>
<td>L</td>
<td>N</td>
<td>AL</td>
</tr>
</tbody>
</table>

Sharada EKAM should be sorted at the top of the script, before ऋ U+11183 SHARADA LETTER A.

2 Description

The Sharada आठ EKAM is an invocation sign used at the beginning of texts. It represents the Sanskrit word पक ekam ‘one’ and is used in place of ekam in salutory phrases, such as om svasti ekam siddham ‘om hail! one, established!’. A detailed description of the character’s origins, meaning, and usage is provided by George A. Grierson in The Journal of the Asiatic Society of Great Britain and Ireland (1916: 678–680), which is reproduced here in figures 1–3.

3 References


4 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.
to the Gurmukhī alphabet, but, unlike them, and like Nāgari, it puts the letters sa and ha at the end of the alphabet, and not after the vowels. Kashmir is called the Śāradā-kṣetra, or holy land of the goddess Śāradā, and this is no doubt the origin of the name of the alphabet, although Elmslie, in his Kashmiri Vocabulary (London, 1872), s.v. Shāradā, mentions a tradition that it is so called in honour of one Śradānandana, who is said to have first reduced the Kāshmirī language to writing.

In India proper, when the alphabet is written down, it is usually preceded by the invocation Ōṁ namah siddhānāṁ, Ōṁ, reverence, established.1 In Kashmir a slightly different formula of invocation is employed, viz. Ōṁ svasti ēkāṁ siddhām, Ōṁ, hail! one, established. As regards the word ēkāṁ, one, it is a curious fact that, while, in writing the invocation, the words Ōṁ, svasti, and siddhām are fully written out—thus, ऋ, स्वस्ति, and सिद्धा—the word ēkāṁ is not written. Instead we have the mystic sign गो, which is named in Kāshmirī oṁ svaṁ gō, and is read as ēkāṁ. So that what is written in the Śāradā character is ऋं स्वस्ति गो गो चिंते read as oṁ svasti ēkāṁ siddhām. The traditional explanation of this is as follows: In order to master the theory of mantras in Kāshmirī Śaivism, it is necessary to learn the meaning or power of each letter composing a mantra, or the mātrā-cakra. Each letter of the alphabet represents some mystic object. The vowels represent the various śūkla, the twenty-five consonants from ka to ma represent the twenty-five lower tattvas, and the other letters the higher tattvas, while kṣa represents the prāṇa-bīja or Life-seed.2 In this way the


2 A full account of the Mātrā-cakra will be found in Kṣemarāja’s Śivasūkraśāra, ii, 7, translated in the Indian Thought Series, No. II.

Figure 1: Description of गो ekam (from Grierson 1916: 678).
letter a represents the ज्ञाना-सक्ति. It also indicates the Supreme (anuttara) and Solitary (akula = kulottirya) Śiva. The sign न is composed of three parts. The horizontal line — represents the letter a, i.e. also Śiva; the two perpendicular strokes ॥ represent the other vowels, and also the saktis; and the two curved marks ◆ represent a plough (hala), and hence all the consonants, which are called by grammarians “hāl”. The whole sign therefore represents all the vowels plus the consonants, or, in other words, the entire alphabet. On the mystical side it also represents Śiva plus all the saktis and tattvas, i.e. Śiva and all his developments in the way of so-called creation.

In the Kāshmirī name ok water, ok means “one”, “non-dual”; sam is a contraction of sāmkītva, or condition of parā samvit, the Supreme Experience; and gor is for gor, it has been inquired into (and therefore understood). With siddham added the whole means “the supreme monist experience has been mastered (for it has been established in the āgamas)”. Eka the siddham has the same meaning.

A less mystical interpretation has been kindly given to me by Professor Barnett. He points out that the siddham is probably derived from the first sūtra of the Kātantra, which runs siddha varṇasaṃmāṇāyaḥ, i.e. “the traditional order of the letters is established (as follows)”, and that this grammar was, over a thousand years ago, the most popular handbook in Northern India and the Buddhist regions of Central Asia. The mark न is evidently one of the sacred symbols used at the commencement or end of any important writing, such as are referred to by Bühler on p. 85 of his Indische Palaeographie, and has practically the force of a sign of punctuation. A not

---

1 When I was in India its use in Northern India seems to have been confined to Eastern Bengal, where I studied it with the local Pāṇḍits. In the rest of Bengal the Mugdhabodha was in general use.
very dissimilar sign will be found at the end of the plate facing p. 281 of vol. ii of Rājendra Lāla Mitra’s Notices of Sanskrit Manuscripts (८८०). Taking this sign as one of punctuation, it would be natural to mark this first division-point by mentally interjecting ेकाम, as a distant echo of the numbering of the first śūtra of the Kātantra, and in course of time the word ेकाम would become petrified, the meaning would be forgotten, and a new mystical meaning given to it.

In those parts of Northern India with which I am acquainted there are, except in the Pañjāb, no special names for the various letters. अ a is called ṝ-कार, े ka is called ka-कāra, and so on. In the Gurmukhi alphabet, used in the Pañjāb, it is different. Here the vowels have each its own name. Thus, initial अ a is called ṝ-र, non-initial उ a is called ṝ-कान, and so on. The consonants are also named by enunciating each twice and doubling the consonant mentioned the second time. Thus झ ka is called kakkā, झ kha is called khakkhā, and so on.

In the schools of Kashmir this Pañjāb system is carried much further. Every vowel form and every consonant has its separate individual name. Most of these names have no definite meaning apart from this connotation, and, as names of letters, do not seem to have been invented on any regular system. Even each syllable of oṃ svastī, and of siddham, and the sign ओ for ेकाम, has its own name. I give these names in the following tables, written in the Nāgarī and Śāradā characters, with a transliteration into the Roman character. As these names are not Sanskrit, but are in the Kāshmirī language, the system of transliteration followed is that which is applied to Kāshmirī, and which differs slightly from the transliteration of the corresponding Nāgarī or Śāradā letters when used for Sanskrit.

Figure 3: Description of ओ ekam (from Grierson 1916: 680).
### A. Administrative

1. **Title:** Proposal to Encode the Sign EKAM for Sharada in ISO/IEC 10646

2. **Requester's name:** Script Encoding Initiative (SEI) / Anshuman Pandey (pandey@umich.edu)

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

4. **Submission date:** 2011-11-02

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

   - **Name of the existing block:** Sharada

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  

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  

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 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?  
   No  

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  

9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?  
   No  

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

11. Does the proposal include use of combining characters and/or use of composite sequences?  
    No  

12. Does the proposal contain characters with any special properties such as control function or similar semantics?  
    No  

13. Does the proposal contain any Ideographic compatibility characters?  
    No  