Proposal to Encode the Siddham Script in ISO/IEC 10646

Anshuman Pandey
Department of History
University of Michigan
Ann Arbor, Michigan, U.S.A.
pandey@umich.edu

July 23, 2012

1 Introduction

Purpose This is a formal proposal to encode the Siddham script in the Universal Character Set (ISO/IEC 10646). It replaces “Preliminary Proposal to Encode Siddham in ISO/IEC 10646” (N4185 L2/12-011R). The proposed encoding is a collaborative effort between the Script Encoding Initiative (SEI) at the University of California, Berkeley and the Shingon Buddhist International Institute, Fresno, California.

Principles of the Encoding Siddham is a Brahmi-based writing system that originated in India, but which at present is used primarily in East Asia. It is associated nowadays with esoteric Buddhist traditions in Japan. The technical description for Siddham given here may differ from the traditional analysis and philosophical interpretations of the script and its constituent characters and glyphs. Siddham is structurally an Indic script and its proposed encoding adheres to the UCS model for Brahmi-based writing systems, such as Devanagari and similar scripts. An attempt has been made to encode all distinct characters attested in Siddham records, although more characters may be uncovered through additional research. The characters that are proposed for encoding have been analyzed in accordance with the character-glyph model of the UCS. As a result, the proposed encoding may contain characters that are not part of traditional character repertoires. It may also exclude characters that are traditionally regarded as independent letters, such as conjuncts, which are to be represented in the manner specified by the UCS encoding model for Indic scripts.

Script Name The script is assigned the name ‘Siddham’. It is also known as ‘Siddhamāṭrkā’ and ‘Kuṭila’, which are specified as aliases in the names list. It has been suggested that ‘Siddhamatrika’ is more suitable than ‘Siddham’ because it broadly accommodates historical and regional forms of the script. However, the the script is most commonly known by the identifier ‘Siddham’. Expert comments on the name are available in L2/12-221 and L2/12-237.

Unification The encoding for Siddham is to serve as a unifying block for all regional variants of the script, such as ‘Siddhamāṭrkā’ and ‘Kuṭila’. The representative glyphs are based upon Japanese forms of Siddham characters on account of active usage of the script by Japanese Buddhist communities.
**Characters Proposed**

A total of 85 Siddham characters is proposed for encoding, as shown in the code chart and names list (figures 1 and 2). Character names follow the UCS convention for Brahmi-based scripts. Other characters have been identified, but are not yet proposed for encoding because additional information regarding their usage is required (see Section 4).

**Allocation**

The Siddham block is allocated to the Supplementary Multilingual Plane (SMP) at the range U+11580..U+115FF.

**Font**

Representative glyphs are based upon the ‘Tenchiji’ font designed by the Shingon Buddhist International Institute with assistance from Lee Collins and Peter Lofting. The present author has added new glyphs and made some modifications to existing glyphs.

## 2 Background

The Siddham script (𑖭 𑖰 𑖾 siddhaṃ, 悉曇文字 xītán wénzì; Jap.悉曇 shittan, Chi.悉曇文字 xītán wénzì; Sans. निद्ध siddham) is a Brahmi-based script historically used for writing Sanskrit in China, Japan, and Korea. It is also known in Japan as bonji (梵字; Chi. fānzì; “Brahmā letters” < Sans. ब्राह्मā brahmā). At present, Siddham is used predominantly in Japan, where it is associated with the Shingon (真言) and Tendai (天台) schools of Japanese esoteric Buddhism (密教 mikkyō). Siddham was used for writing Buddhist manuscripts, but it is now mainly used for ceremonial and ritualistic purposes, such as the copying of sūtra-s and the writing of mantra-s and bijāksara-s “seed syllables” associated with esoteric Buddhist practices. It was also used for inscriptions on statuary, amulets, and other artefacts.

Tradition states that the esoteric philosophy for Siddham was introduced in Japan in 806 by the Buddhist monk Kūkai (空海) (774–835), who learned the script in China from Prajñā (Sans. प्रज्ञा; Chi. 般若三藏 Bōrě Sāncáng) (734–c.810), a monk from Kashmir who served the Buddhist community of Chang’an (modern Xi’an) as an expert of Sanskrit and translator of Buddhist texts (van Gulik 1980: 114; Abe 1999: 117–119). Known better by his posthumous title Kōbō-Daishi (弘法大師), Kūkai was the founder of the Shingon (“True Word”) school. Another proponent of Siddham in Japan was Saichō (最澄) (767–822), also called Dengyō-Daishi (傳教大師), who was the founder of the Tendai school.

There are several regional forms of Siddham. The Japanese form is described here and it is presented as the normative form for the encoding. In Japan, Siddham calligraphy is classified as ‘formal’ or ‘informal’ based upon the implement used for writing. The ‘formal’ script is written using a wooden stylus called bokuhitsu (木筆; Chi. mù-bǐ), while the ‘informal’ or ‘brush’ style is written using the common hair-tipped fude (筆; Chi. 毛筆 máo-bǐ). The ‘formal’ and ‘informal’ styles of letters are shown in figures 12–17. The forms of Siddham letters used in Korea (see figures 21–26) differ significantly from those of Japanese and Chinese Siddham. However, given the close relationships between the regional forms and their identities as ‘Siddham’, it is practical to unify these variants with the proposed script block.

Siddham is most closely related to Sharada (see tables 1 and 2), a Brahmi-based script that originated in Kashmir and that was used throughout northwestern India for producing manuscripts and inscriptions (see Pandey 2009). The use of Sharada spread from India into Central Asia and travelled from there to China with Buddhist monks. Although there are not many records in Central Asian Siddham, an example of it is found on the palm-leaf manuscripts of the Hōryū-ji (法隆寺) temple in Nara, Japan (see figure 6), which were brought there from China in 609. Unlike Siddham written with the bokuhitsu or fude, the Siddham letters on the Hōryū-ji manuscripts were written with the typical Indian reed-pen. A detailed palaeographical description of the Siddham of these manuscripts was presented by Georg Bühler (1884).
3  Writing System

3.1  Structure

The general structure (phonetic order, mātrā reordering, use of virāma, etc.) of Siddham is similar to that of Devanagari.

3.2  Directionality

Siddham is written horizontally left-to-right and also vertically top-to-bottom with lines proceeding from right-to-left. See figure 5 for an example of text written top-to-bottom.

3.3  Virāma

The Siddham sign 𑖘 is identical to the corresponding character in Devanagari. It is used for silencing the inherent vowel of a consonant. The default rendering of virāma is as a visible sign. The virāma is known in Sanskrit as हल᭠त halanta and in Japanese as 怛達点 tatatsu-ten.

3.4  Vowel Letters

Fourteen vowel letters are proposed for encoding:

<table>
<thead>
<tr>
<th>औ</th>
<th>उ</th>
<th>व</th>
<th>ओ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ा</td>
<td>ा</td>
<td>ा</td>
<td>ा</td>
</tr>
<tr>
<td>ि</td>
<td>ि</td>
<td>ि</td>
<td>ि</td>
</tr>
</tbody>
</table>

Encoding Order  The encoding order for vowel signs follows that of Brahmi-based scripts in the UCS. In the Japanese arrangement for Siddham, the vocalic letters appear at the end of the vowel order.

Graphical variants  Variant forms of vowel letters are attested, eg. ा is also written as  (see figure 12). These are to be managed through fonts.

3.5  Vowel Signs

Eleven dependent vowel signs are proposed for encoding:

<table>
<thead>
<tr>
<th>𑖘 VOWEL SIGN AA</th>
<th>𑖘 VOWEL SIGN UU</th>
<th>𑖘 VOWEL SIGN AI</th>
</tr>
</thead>
<tbody>
<tr>
<td>𑖘 VOWEL SIGN I</td>
<td>𑖘 VOWEL SIGN VOCALIC R</td>
<td>𑖘 VOWEL SIGN O</td>
</tr>
<tr>
<td>𑖘 VOWEL SIGN II</td>
<td>𑖘 VOWEL SIGN VOCALIC RR</td>
<td>𑖘 VOWEL SIGN AU</td>
</tr>
<tr>
<td>𑖘 VOWEL SIGN U</td>
<td>𑖘 VOWEL SIGN E</td>
<td>𑖘 VOWEL SIGN O</td>
</tr>
</tbody>
</table>
Unattested vowel signs  Dependent forms of 𑖈 VOWEL LETTER VOCALIC L and 𑖇 VOWEL LETTER VOCALIC LL are as of yet unattested. Space has been reserved in the code chart in the event that additional research uncovers such characters.

Encoding Order  The encoding order for vowel signs follows that of Brahmi-based scripts in the UCS. In the Japanese arrangement for Siddham, the signs for the vocalic letters appear at the end of the vowel-sign order.

Graphical variants  There are glyphic variants of vowel signs (see figures 12 and 19), which are to be managed through fonts. One of the most common variants is the use of 𑖈 VOWEL SIGN AI for 𑖇.

Alternate forms  There is an alternate form for 𑖉 VOWEL SIGN VOCALIC RR. Normally, this sign is written as 𑖓. The two-part form 𑖏 is also attested; it is a combination of 𑖇 VOWEL SIGN VOCALIC R and 𑖕 VOWEL SIGN AA, in which the latter sign acts as a vowel-length mark.

3.6 Vowel-Modifier Signs

Three vowel-modifier signs are proposed for encoding:

1. 𑖏 SIGN ANUSVARA  This sign indicates nasalization.

2. 𑖎 SIGN CANDRABINDU  This sign indicates nasalization.

3. 𑖐 SIGN VISARGA  This sign represents post-vocalic aspiration (/h/). In Sanskrit phonology, it indicates an allophone of 𑖉 SA or 𑖊 RA in word-final position.

Ordering of multiple signs  In certain cases, such as when writing ‘seed syllables’, a base letter that is combined with the signs ANUSVARA or CANDRABINDU may also be combined with the sign VISARGA, eg. 𑖋 or 𑖎 𑖐 𑖎. In such cases the VISARGA is written last: <𑖋 AA, 𑖎 CANDRABINDU, 𑖐 VISARGA> → 𑖋 𑖐 𑖎. Rendering engines should recognize such sequences of combining marks as valid.

3.7 Consonants

Thirty-three consonant letters are proposed for encoding:

| 𑖉 | KA  |
| 𑖊 | KHA |
| 𑖋 | GA  |
| 𑖌 | GHA |
| 𑖍 | NGA |
| 𑖎 | CA  |
| 𑖏 | CHA |
| 𑖐 | JA  |
| 𑖑 | JHA |

| 𑖙 | NYA |
| 𑖚 | TTA |
| 𑖜 | TTHA |
| 𑖝 | DDA |
| 𑖞 | DDHA |
| 𑖟 | NNA |
| 𑖠 | TA  |
| 𑖡 | THA |
| 𑖢 | DA  |

| 𑖣 | DHA |
| 𑖤 | NA  |
| 𑖥 | PA  |
| 𑖦 | PHA |
| 𑖧 | BA  |
| 𑖨 | BHA |
| 𑖩 | MA  |
| 𑖪 | YA  |
| 𑖫 | RA  |

| 𑖬 | LA  |
| 𑖭 | VA  |
| 𑖮 | SHA |
| 𑖯 | SSA |
| 𑖰 | SA  |
| 𑖱 | HA  |

| 𑖲 | LAA |
| 𑖳 | VAA |
| 𑖴 | SHA |
| 𑖵 | SSA |
| 𑖶 | SA  |
| 𑖷 | HA  |
Each consonant bears the inherent vowel /a/, represented by र A, which is silenced using ओ VIRAMA. Variant forms of consonants are attested (see figures 13–17); these are to be managed at the font level.

3.8 NUKTA

The sign ओ NUKTA is used for transcribing sounds for which distinct characters do not natively exist in the writing system (see figure 43). The NUKTA is not a traditional Siddham character, but it is part of 現代悉曇 gendai shittan “modern Siddham”, which is an extension of Siddham that accommodates the writing of Japanese and English. The name of the character is derived from the Arabic word نُقْطَة nuqtah = nukta “dot” and corresponds to characters such as ओ U+093C DEVANAGARI SIGN NUKTA.

The NUKTA is generally written with a letter that has the closest phonetic proximity to the target sound. For example, in figure 44 it is shown combined with ए JA (/ʤ/) for representing /z/. The NUKTA is always combined with the letter that it modifies. When it occurs with a consonant that is also combined with a vowel sign, then NUKTA is written immediately after the consonant, eg. <ए JA, ओ SIGN NUKTA, ओ VOWEL SIGN U> → एँ zu. Its rendering depends upon various typographic considerations.

3.9 Consonant Conjuncts

Consonant clusters in Siddham are written as conjuncts and follow the same model as conjuncts in Devanagari. Conjuncts are represented using virama, which is written after each consonant in a cluster. They may be written vertically, horizontally, and as independent ligatures.

There are traditional Chinese and Japanese tabulations for Siddham conjuncts. One such classification system was described in the manual Xītán Zìjì (悉曇字記) by the Chinese monk Zhì-guǎng (智廣) (d. 806), who organized the various combinations of Siddham consonant clusters into eighteen categories (see figure 45). Examples of several conjuncts are shown in figures 46–61.

**Shaping**  The shapes of some consonants change when they occur in conjunct-initial position (eg. ए*K*A → एः; ऐ RA → ऐ repha), some are written using post-base forms when they are non-initial (eg. ए YA → ए). These ligating forms are illustrated in figures 13–17 (column labeled “in combination”) and figures 19–20.

**Depth**  Siddham conjuncts may represent clusters with large numbers of consonants. For example, a conjunct of six consonants is एः*k*K*A, which is produced using the sequence <़ RA, ओ VIRAMA, ए* K*A, ओ VIRAMA, ए* SSA, ओ VIRAMA, ए* VA, ओ VIRAMA, ए* RA, ओ VIRAMA, ए* YA>.

**Special Conjuncts**  The glyphs एः ksa and एः llam are often shown as independent letters in Siddham charts. The ksa represents conjuncts of different consonant letters and llam represents conjuncts of geminate consonants. Based upon the UCS encoding model, they are not independent letters, but conjuncts. They are to be represented using the following sequences: एः = <़ KA, ओ VIRAMA, ए* SSA> (compare Devanagari एः ksa = <़ KA, ओ VIRAMA, ए* SSA>); एः = <़ LA, ओ VIRAMA, ए* LA, ओ ANUSVARA>.

**Default Representation**  It is expected that a basic Siddham font will contain the necessary glyphs for properly representing conjuncts. A consonant cluster is depicted with the appropriate conjunct glyph only if such a glyph is available in the font. If the conjunct glyph is unavailable, the bare consonants in the cluster are depicted using their full forms combined with a visible virama. While these depictions of conjuncts may not be used in Siddham orthography, they are necessary for the representation of the script in plain text. In order to render Siddham text appropriately in the absence of distinct glyphs for any possible conjunct, it is
recommended that a set of condensed forms of letters in conjunct-initial and non-initial positions be available so that arbitrary conjuncts may be rendered in the vertical orientation used in Siddham orthography.

### 3.10 Consonant-Vowel Combinations

Vowel signs combine with consonant letters and conjuncts as in Devanagari. Some consonant shapes are modified when they are written with vowels (e.g., 𑖌 KA → 𑖌 when combining with 𑖗 VOWEL SIGN U: 𑖙 ku). Some vowel signs are written as contextual forms when they occur with certain consonants. For example, 𑖘 VOWEL SIGN U → 𑖘 and 𑖘 VOWEL SIGN UU → 𑖘 when combined with 𑖙 NG, 𑖙 JA, 𑖙 TTA, 𑖙 TTHA, 𑖙 DDHA, 𑖙 NA, 𑖙 PA, 𑖙 VA. These signs have specific forms when written with 𑖙 RA: 𑖙 ru and 𑖙 rū.

### 3.11 Invocations

One invocation sign is proposed for encoding:

1. 𑗁 SIGN SIDDHAM  This sign is written at the beginning of a text (see figures 6 and 38). Palaeographically, the sign corresponds to characters used in other scripts, such as 𑖙 U+0FD3 TIBETAN MARK INITIAL BRDA RNYYING YIG MGO MDUN MA. It represents the Sanskrit word शिरम siddham “accomplished” and the phrase शिरमस्य siddhirastu “may there be success”. The sign is often glossed as ‘ōm’ (see figure 27). There is a vertically-oriented glyphic variant: 𑗁 (see figure 33). In the Japanese tradition, SIDDHAM is analyzed as being formed from the lower portion of 𑖙 t.

The syllable om is not proposed for independent encoding. It is to be represented using the sequence <𑖙 O, 𑖙 CANDRABINDU> → 𑖙. A distinct om character is not attested for Siddham.

### 3.12 Punctuation

Four punctuation characters are proposed for encoding:

1. ꞓ DANDA  This mark is used for marking the end of sentences and other short text sections (see figures 32, 40, 41). It has a graphical variant ꞓ that is used in ‘informal’ Japanese Siddham. The DANDA corresponds to, but is graphically distinct from, ꞓ U+0964 DEVANAGARI DANDA.

2. ꞔ DOUBLE DANDA  This mark is used at the end of paragraphs and larger text blocks (see figures 32, 40, 41). It is also written with an underdot ꞔ. The graphical variant ꞔ is used in ‘informal’ Japanese writing. The DOUBLE DANDA corresponds to, but is graphically distinct from, ꞔ U+0965 DEVANAGARI DOUBLE DANDA.

3. ꞓ WORD SEPARATOR ONE  This sign is used for marking boundaries between words. It is written at the head-height and is distinct from the ꞓ middle dot. Usage of WORD SEPARATOR ONE is shown in figure 38.

4. ꞔ WORD SEPARATOR TWO  This sign is used for marking boundaries between words and phrases. Its usage is shown in figures 40, 41, 42, where it appears alongside ꞓ DANDA and ꞔ DOUBLE DANDA.

The following punctuation mark is not proposed for separate encoding:

1. ꞔ  Word and phrase separation is also indicated using a ꞔ middle dot (see figure 39). This mark differs from ꞔ WORD SEPARATOR ONE in terms of its vertical placement. It is to be written using the already encoded ꞔ U+00B7 MIDDLE DOT.
3.13 Section Marks

The following 14 characters are used for marking the end of sections: 𑗉 SECTION MARK ONE, 𑗊 SECTION MARK TWO, 𑗋 SECTION MARK THREE, 𑗌 SECTION MARK FOUR, 𑗍 SECTION MARK FIVE, 𑗎 SECTION MARK SIX, 𑗏 SECTION MARK SEVEN, 𑗐 SECTION MARK EIGHT, 𑗑 SECTION MARK NINE, 𑗒 SECTION MARK TEN, 𑗓 SECTION MARK ELEVEN, 𑗔 SECTION MARK TWELVE, 𑗕 SECTION MARK THIRTEEN, 𑗖 SECTION MARK FOURTEEN.

The first twelve marks are attested in various Siddham manuals (see figures 32, 33, 35). The last two are attested in a manuscript (see figure 39). The 𑗔 SECTION MARK TWELVE is graphically distinct from other section marks, perhaps also functionally, and could be encoded as a completion-of-text mark.

3.14 Repetition Marks

The 𑖆 REpetition MARK ONE, 𑖇 REpetition MARK TWO, and 𑖈 REpetition MARK THREE are used for indicating the repetition of text. They are written after the text that is to be repeated.

The 𑖆 REpetition MARK ONE has the glyphic variants 𑖅 and 𑖆 (see figures 32, 33, 35). Based upon its graphical shape and function, it is likely that REpetition MARK ONE is derived from an Indic digit ‘2’ (compare Devanagari २). A common practice in manuscript traditions of northern India is to indicate the repetition of words using digits. The digit used indicates the number of repetitions, eg. ‘2’ for twice, ‘3’ for thrice. As there are no attested digits for Siddham, it appears that the original link between REpetition MARK ONE and a numeric value was forgotten and the original digit for ‘2’ was preserved as a symbol. This explains the serialized usage of REpetition MARK ONE in the text excerpt in figure 36 for indicating a triple reading, ie. Devanagari “ᱠं २ २” instead of “ᱠं ३” for hūṃ hūṃ hūṃ).

The 𑖇 REpetition MARK TWO is used in the vertical version of the Uṣṇiṣa Vijaya Dhāraṇī Sūtra shown in figure 5. An excerpt of the Sūtra with the repetition marks highlighted is shown in figure 37. Here, the mark appears after 𑖈 ya (3rd character), another ya (7th character), and 𑖉 ra (10th character). The mark prompts the reader to parse the text as “jaya jaya, vijaya vijaya, smara smara”. It is possible that REpetition MARK TWO is used only in a vertical environment. This is supported by the accompanying Japanese annotation, which glosses 𑖇 using ｶﾝｶﾞﾛｳ Vertical Kana REPEAT MARK.

The 𑖈 REpetition MARK THREE is shown in several Siddham manuals. It is proposed for encoding because it is a graphically-distinct character.

4 Characters Not Proposed for Encoding

Several characters attested in Siddham sources are not presently proposed for encoding. Additional research is needed for determining the suitability of encoding these characters, which are described below.

4.1 Conjuncts

The glyphs 𑖉 kṣa and 𑖊 llaṃ are conjuncts and are to be represented as such (see Section 3.9).

4.2 Digits

There is no traditional set of Siddham digits. There are, however, modern innovations named “Bonji numerals” (see figure 18). These digits require additional research and may be considered for future inclusion.
4.3 Editorial Marks

Siddham manuscripts contain editorial marks, such as those used for indicating the insertion or deletion of text. The characters \( \text{•} \) and \( \text{•} \) are used for marking deletions (see figure 35). Another character is a caret-like mark that is written below the baseline to indicate the insertion of text. These editorial marks are not unique to Siddham and are used in various Indic manuscript traditions. These characters should be represented using already-encoded characters with similar functions.

4.4 Pedagogical Characters

There are several pedagogical characters used for describing Siddham letters in script manuals. Some of these are discussed below. These characters may be candidates for encoding, but additional research is required regarding their usage and suitability for encoding, as per the UCS encoding model.

Character Strokes Siddham manuals describe elemental character strokes that are used for writing letters (see figure 65). Some of these are described below.

1. \( \text{•} \) myō-ten The stroke \( \text{•} \) is the initial mark made when a brush is pressed upon a surface for writing any Siddham letter (see the stroke order in figures 66 and 67). It is known in Japanese as 命点 myō-ten “life mark”. The myō-ten corresponds to the Chinese simple stroke \( \text{•} \) diǎn “dot”, which is encoded in the UCS as \( \text{•} \) \( \text{•} \)          .

2. \( \text{•} \) a-ten The stroke \( \text{•} \) is shown in some historical and modern Siddham handbooks as a dependent vowel sign of \( \text{•} \) (see figures 19, 63, 64) or as a ‘variation’ of the letter (see figure 12). For this reason it is called ア点 a-ten “a mark” in Japanese. It is a horizontally elongated form of \( \text{•} \) myō-ten. The a-ten is not a true ‘vowel sign’, but a portion of the top-stroke of each consonant letter.

3. \( \text{•} \) gyōgatsu-ten The stroke \( \text{•} \) is known in Japanese as 仰月点 gyōgatsu-ten “moon-viewing mark” and in Sanskrit as अधच奄 ardhacandra “half-moon”. It is the foundational stroke of the sign \( \text{•} \) CANDRABINDU, which is produced by writing \( \text{•} \) ANUSVARA with gyōgatsu-ten. The stroke is considered 莊厳 shougon (Sans. अलंकार aḷaṃkāra) “adornment” or “decoration” in that it does not change the phonological value of ANUSVARA. The use of gyōgatsu-ten is shown in figure 68, where it appears as an independent graphical element.

Figure 13 (row 11) shows the below-base character \( \text{•} \) listed along with \( \text{•} \) CANDRABINDU and \( \text{•} \) ANUSVARA. What is this character? Is this the gyōgatsu-ten?

4. \( \text{•} \) u-ten The stroke \( \text{•} \) is known as ウ点 u-ten “u mark” in Japanese. It is used for forming characters like \( \text{•} \) 1, \( \text{•} \) U, and \( \text{•} \) U.

5. \( \text{•} \) en-ten The stroke \( \text{•} \) is known as 円点 en-ten “circle mark” in Japanese. It is used for forming characters like \( \text{•} \) ANUSVARA and \( \text{•} \) TTHA.

‘Headless’ Letters Some Japanese treatises on Siddham offer explanations about the phonological properties of Siddham consonants by using ‘headless’ letters, which are the regular consonants written without the top-stroke (myō-ten; Sans. mātrā). One such work, the Shittan Bunsho (悉曇聞書) of Jiun Sonja (慈雲尊者) (1718–1804), expounds upon the alphasyllabic and alphabetic natures of Siddham consonants by using ‘headless’ letters. Jiun explains that removing the top-stroke of थ ka produces the ‘headless’ form घ in which the inherent vowel is absent: /ka/ \( \rightarrow \) /k/. According to this view, the inherent vowel /a/ is contained in the top-stroke and eliminating it produces an alphabetic form of a letter. Secondly, when this ‘headless’ ka is combined with a \( \text{•} \) VIRAMA as घ, the remaining consonantal value is removed: /k/ \( \rightarrow \) Ø. In
this case, the virama essentially eradicates the entire letter. A table of ‘headless’ letters is given by Jōgon (浄嚴) (1639–1702) in his Shittan Sanmitsushō (浄厳 三密鈔), which is presented here in figure 62. While ‘headless’ forms of Siddham consonants are required for accurately representing Jiun’s Shittan Bunsho, they are to be considered as glyphic variants of regular letters and are to be managed through fonts.

4.5 Bījāksara-

A bījāksara or bīja “seed syllable” may be written such that the constituent characters are joined to form a highly-stylized ligature (see figures 70, 71, 72, 73). For example, the word ṇ phat (< pha, tta, ạ virama>) may be written as ṇ phat. A more complex example is a bīja of Acala (Sans. अचल; Chi. 不動明王 BúdòngMíngwáng; Jap. FudōMyōō), which is a ligature formed by joining ṇ hāṃ and ṇ māṃ, the last two syllables of his mantra, into a monogram written as ṇ or ṇ hāṃmāṃ.

These types of ligatures cannot be represented in plain text using characters proposed for encoding. They are calligraphic innovations and are to be managed through the application of font features or text attributes.

5 Character Data

5.1 Collation

The primary collating order for Siddham is as follows:

A A < AA < I < II < U < UU < E < AI < O < AU <
 exacerbated vocalic R < exacerbated vocalic RR < vocalic L < vocalic LL < KA < KHA <
 GA < GHA < NGA < CA < CHA < JA < JHA < NYA <
 TTA < TTHA < DDA < DDHA < NNA < TA < THA < DA <
 DHA < NA < PA < PHA < BA < BHA < MA < YA < RA <
 LA < VA < SHA < SSA < SA < HA < SIGN AA < SIGN I <
 SIGN II < SIGN U < SIGN UU < SIGN E < SIGN AI < SIGN O <
 SIGN AU < SIGN VOCALIC R < SIGN VOCALIC RR < SIGN VIRAMA

The following signs have secondary weights: _CLEARBINDU, pageNum}.
5.3 Linebreaking

Linebreaking properties given in the format of LineBreak.txt:

<table>
<thead>
<tr>
<th>Range</th>
<th>Description</th>
<th>Characters</th>
</tr>
</thead>
<tbody>
<tr>
<td>11580..115AE</td>
<td>AL # LETTER A .. LETTER HA</td>
<td>14</td>
</tr>
<tr>
<td>115AF..115BF</td>
<td>CM # SIGN AA .. SIGN VISARGA</td>
<td>2</td>
</tr>
<tr>
<td>115C0</td>
<td>BB # SIGN SIDDHAM</td>
<td>1</td>
</tr>
<tr>
<td>115C1..115C4</td>
<td>BA # DANDA .. WORD SEPARATOR TWO</td>
<td>4</td>
</tr>
<tr>
<td>115C5..115C7</td>
<td>AL # REPETITION MARK ONE .. REPETITION MARK THREE</td>
<td>3</td>
</tr>
<tr>
<td>115C8..115D5</td>
<td>BA # SECTION MARK ONE .. SECTION MARK FOURTEEN</td>
<td>8</td>
</tr>
</tbody>
</table>

5.4 Syllabic Categories

Syllabic categories given in the format of IndicSyllabicCategory.txt:

<table>
<thead>
<tr>
<th>Range</th>
<th>Description</th>
<th>Characters</th>
</tr>
</thead>
<tbody>
<tr>
<td># Indic_Syllabic_Category=Bindu</td>
<td>115BD..115BE ; Bindu # Mn [2] SIGN CANDRABINDU .. SIGN ANUSVARA</td>
<td>14</td>
</tr>
<tr>
<td># Indic_Syllabic_Category=Visarga</td>
<td>115BF ; Visarga # Mc SIGN VISARGA</td>
<td>2</td>
</tr>
<tr>
<td># Indic_Syllabic_Category=Virama</td>
<td>115BC ; Virama # Mn SIGN VIRAMA</td>
<td>2</td>
</tr>
<tr>
<td># Indic_Syllabic_Category=Nukta</td>
<td>115C0 ; Nukta # Mn SIGN NUKTA</td>
<td>2</td>
</tr>
<tr>
<td># Indic_Syllabic_Category=Vowel_Independent</td>
<td>11580..1158D ; Vowel_Independent # Lo [14] LETTER A .. LETTER AU</td>
<td>14</td>
</tr>
<tr>
<td># Indic_Syllabic_Category=Vowel_Dependent</td>
<td>115AF..115B1 ; Vowel_Dependent # Mc [3] VOWEL SIGN AA .. VOWEL SIGN II</td>
<td>12</td>
</tr>
<tr>
<td># Indic_Syllabic_Category=Vowel_Dependent</td>
<td>115B2..115B3 ; Vowel_Dependent # Mn [2] VOWEL SIGN U .. VOWEL SIGN UU</td>
<td>2</td>
</tr>
<tr>
<td># Indic_Syllabic_Category=Vowel_Dependent</td>
<td>115B4..115B5 ; Vowel_Dependent # Mn [2] VOWEL SIGN VOCALIC R .. VOWEL SIGN VOCALIC RR</td>
<td>2</td>
</tr>
<tr>
<td># Indic_Syllabic_Category=Vowel_Dependent</td>
<td>115B8 ; Vowel_Dependent # Mc VOWEL SIGN E</td>
<td>1</td>
</tr>
<tr>
<td># Indic_Syllabic_Category=Vowel_Dependent</td>
<td>115B9 ; Vowel_Dependent # Mn VOWEL SIGN AI</td>
<td>1</td>
</tr>
<tr>
<td># Indic_Syllabic_Category=Vowel_Dependent</td>
<td>115BA..115BB ; Vowel_Dependent # Mc [2] VOWEL SIGN O .. VOWEL SIGN AU</td>
<td>2</td>
</tr>
<tr>
<td># Indic_Syllabic_Category=Consonant</td>
<td>11599..1159F ; Consonant # Mc CONSONANT</td>
<td>10</td>
</tr>
</tbody>
</table>
1158E..115AE ; Consonant # Lo [33] LETTER KA .. LETTER HA

5.5 Matra Categories

Matra categories given in the format of IndicMatraCategory.txt:

# Indic_Matra_Category=Right
115AF ; Right # Mc VOWEL SIGN AA
115B1 ; Right # Mc VOWEL SIGN II

# Indic_Matra_Category=Left
115B0 ; Left # Mc VOWEL SIGN I
115B8 ; Left # Mc VOWEL SIGN E

# Indic_Matra_Category=Left_And_Right
115BA ; Left_And_Right # Mc VOWEL SIGN O

# Indic_Matra_Category=Top
115B9 ; Top # Mn VOWEL SIGN AI

# Indic_Matra_Category=Top_And_Left_And_Right
115BB ; Top_And_Left_And_Right # Mc VOWEL SIGN AU

# Indic_Matra_Category=Bottom
115B2..115B3 ; Bottom # Mn [2] VOWEL SIGN U .. VOWEL SIGN UU
115B4..115B5 ; Bottom # Mn [2] VOWEL SIGN VOCALIC R .. VOWEL SIGN VOCALIC RR
115BC ; Bottom # Mn SIGN VIRAMA

6 References


Proposal to Encode the Siddham Script in ISO/IEC 10646
Anshuman Pandey


7 Acknowledgments

I am thankful to Lee Collins (Apple Computer, Inc.) for reviewing drafts of this proposal, for offering background and usage information on Siddham, and for providing samples of Korean Siddham, as well as the ‘Tenchiji’ font used here. Equally important are the contributions of Eijun Eidson of the Shingon Buddhist International Institute, who is headmaster of the Koyasan Shingon Temple of North America. I am grateful to Shriramana Sharma for offering corrections on technical matters regarding the representation of virama and alternate forms of vowel signs; also to Vinodh Rajan for identifying the use of nukta in ‘modern’ Siddham. I am indebted to Micah Auerback (Asian Languages and Cultures, University of Michigan, Ann Arbor) and Wang Xiyue (王夕越) for patiently answering my questions about Japanese and Chinese sources, respectively, and for providing insight into texts through their own expert translations. Stefan Baums (University of California, Berkeley), Andrew Glass (Microsoft), Kengo Harimoto (Nepal-German Manuscript Cataloguing Project), Vinodh Rajan, Shriramana Sharma, and Iain Sinclair (Monash University) provided comments on the name for the script block.

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.
<table>
<thead>
<tr>
<th>U+1158</th>
<th>U+1159</th>
<th>U+115A</th>
<th>U+115B</th>
<th>U+115C</th>
<th>U+115D</th>
<th>U+115E</th>
<th>U+115F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Proposed code chart for Siddham.
The script is also known by the names 'Siddhamatrika' and 'Kutila'.

Independent vowels

11580 𑖀 SIDDHAM LETTER A
11581 𑖁 SIDDHAM LETTER AA
11582 𑖂 SIDDHAM LETTER I
11583 𑖃 SIDDHAM LETTER II
11584 𑖄 SIDDHAM LETTER U
11585 𑖅 SIDDHAM LETTER UU
11586 𑖆 SIDDHAM LETTER VOCALIC R
11587 𑖇 SIDDHAM LETTER VOCALIC RR
11588 𑖈 SIDDHAM LETTER VOCALIC L
11589 𑖉 SIDDHAM LETTER VOCALIC LL
1158A 𑖊 SIDDHAM LETTER E
1158B 𑖋 SIDDHAM LETTER AI
1158C 𑖌 SIDDHAM LETTER O
1158D 𑖍 SIDDHAM LETTER AU

Consonants

11590 𑖎 SIDDHAM LETTER KA
11591 𑖏 SIDDHAM LETTER KHA
11592 𑖐 SIDDHAM LETTER GA
11593 𑖑 SIDDHAM LETTER NGA
11594 𑖒 SIDDHAM LETTER CA
11595 𑖓 SIDDHAM LETTER CHA
11596 𑖔 SIDDHAM LETTER JA
11597 𑖕 SIDDHAM LETTER JHA
11598 𑖖 SIDDHAM LETTER NYA
11599 𑖗 SIDDHAM LETTER TTA
1159A 𑖘 SIDDHAM LETTER TTHA
1159B 𑖙 SIDDHAM LETTER DDA
1159C 𑖚 SIDDHAM LETTER DDHA
1159D 𑖛 SIDDHAM LETTER NNA
1159E 𑖜 SIDDHAM LETTER NGA
1159F 𑖝 SIDDHAM LETTER NNYA
115A0 𑖞 SIDDHAM LETTER DHA
115A1 𑖟 SIDDHAM LETTER DA
115A2 𑖠 SIDDHAM LETTER NA
115A3 𑖡 SIDDHAM LETTER PA
115A4 𑖢 SIDDHAM LETTER PHA
115A5 𑖣 SIDDHAM LETTER BA
115A6 𑖤 SIDDHAM LETTER BHA
115A7 𑖥 SIDDHAM LETTER MA
115A8 𑖦 SIDDHAM LETTER YA
115A9 𑖧 SIDDHAM LETTER RA
115AA 𑖨 SIDDHAM LETTER LA
115AB 𑖩 SIDDHAM LETTER VA
115AC 𑖪 SIDDHAM LETTER SHA
115AD 𑖫 SIDDHAM LETTER SSA
115AE 𑖬 SIDDHAM LETTER SA
115AF 𑖭 SIDDHAM LETTER HA

Dependent vowel signs

115B0 𑖮 SIDDHAM VOWEL SIGN O
115B1 𑖯 SIDDHAM VOWEL SIGN AU
115B2 𑖰 SIDDHAM VOWEL SIGN VOCALIC R
115B3 𑖱 SIDDHAM VOWEL SIGN VOCALIC RR
115B4 𑖲 SIDDHAM VOWEL SIGN VOCALIC L
115B5 𑖳 SIDDHAM VOWEL SIGN VOCALIC LL
115B6 𑖴 SIDDHAM VOWEL SIGN VOCALIC U
115B7 𑖵 SIDDHAM VOWEL SIGN VOCALIC UU
115B8 𑖶 SIDDHAM VOWEL SIGN VOCALIC II
115B9 𑖷 SIDDHAM VOWEL SIGN VOCALIC I
115BA 𑖸 SIDDHAM VOWEL SIGN VOCALIC AA
115BB 𑖹 SIDDHAM VOWEL SIGN VOCALIC E
115BC 𑖺 SIDDHAM VOWEL SIGN VOCALIC AI
115BD 𑖻 SIDDHAM VOWEL SIGN VOCALIC EE
115BE 𑖼 SIDDHAM VOWEL SIGN VOCALIC AE
115BF 𑖽 SIDDHAM VOWEL SIGN VOCALIC AA

Various signs

115C0 𑖾 SIDDHAM SIGN NUKTA
115C1 𑖿 SIDDHAM SIGN VISARGA
115C2 𑖺 SIDDHAM SIGN ANUSVARA
115C3 𑖻 SIDDHAM SIGN CANDRABINDU
115C4 𑖼 SIDDHAM SIGN KUTILA
115C5 𑖽 SIDDHAM SIGN VIRAMA

Invocation sign

115C6 𑖾 SIDDHAM SIGN SIDDHAM

Punctuation

115C7 𑖿 SIDDHAM WORD SEPARATOR ONE
115C8 𑖽 SIDDHAM WORD SEPARATOR TWO
115C9 𑖾 SIDDHAM SECTION MARK ONE
115CA 𑖿 SIDDHAM SECTION MARK TWO
115CB 𑖽 SIDDHAM SECTION MARK THREE
115CC 𑖾 SIDDHAM SECTION MARK FOUR
115CD 𑖿 SIDDHAM SECTION MARK FIVE
115CE 𑖽 SIDDHAM SECTION MARK SIX
115CF 𑖾 SIDDHAM SECTION MARK SEVEN
115D0 𑖿 SIDDHAM SECTION MARK EIGHT
115D1 𑖽 SIDDHAM SECTION MARK NINE
115D2 𑖾 SIDDHAM SECTION MARK TEN
115D3 𑖿 SIDDHAM SECTION MARK ELEVEN
115D4 𑖽 SIDDHAM SECTION MARK TWELVE
115D5 𑖾 SIDDHAM SECTION MARK THIRTEEN
115D6 𑖿 SIDDHAM SECTION MARK FOURTEEN

Section marks

These characters are used for indicating the end of text

115D7 𑖽 SIDDHAM SECTION MARK SEVEN
115D8 𑖾 SIDDHAM SECTION MARK EIGHT
115D9 𑖿 SIDDHAM SECTION MARK NINE
115DA 𑖽 SIDDHAM SECTION MARK TEN
115DB 𑖾 SIDDHAM SECTION MARK ELEVEN
115DC 𑖿 SIDDHAM SECTION MARK TWELVE
115DD 𑖽 SIDDHAM SECTION MARK THIRTEEN
115DE 𑖾 SIDDHAM SECTION MARK FOURTEEN

Figure 2: Proposed names list for Siddham.
<table>
<thead>
<tr>
<th>Siddham</th>
<th>Sharada</th>
<th>Nagari</th>
<th>Siddham</th>
<th>Sharada</th>
<th>Nagari</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>अ</td>
<td>अ</td>
<td>-A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AA</td>
<td>आ</td>
<td>आ</td>
<td>-AA</td>
<td>ा</td>
<td>ा</td>
</tr>
<tr>
<td>I</td>
<td>ई</td>
<td>ई</td>
<td>-I</td>
<td>नि</td>
<td>नि</td>
</tr>
<tr>
<td>II</td>
<td>उ</td>
<td>उ</td>
<td>-II</td>
<td>ऊ ऊ</td>
<td>ऊ ऊ</td>
</tr>
<tr>
<td>U</td>
<td>ऊ</td>
<td>ऊ</td>
<td>-U</td>
<td>ऋ ऋ</td>
<td>ऋ ऋ</td>
</tr>
<tr>
<td>UU</td>
<td>ऋ</td>
<td>ऋ</td>
<td>-UU</td>
<td>ओ ओ</td>
<td>ओ ओ</td>
</tr>
<tr>
<td>R</td>
<td>र</td>
<td>र</td>
<td>-R</td>
<td>र र</td>
<td>र र</td>
</tr>
<tr>
<td>RR</td>
<td>र</td>
<td>र</td>
<td>-RR</td>
<td>र र</td>
<td>र र</td>
</tr>
<tr>
<td>L</td>
<td>ल</td>
<td>ल</td>
<td>-L</td>
<td>ल ल</td>
<td>ल ल</td>
</tr>
<tr>
<td>LL</td>
<td>ल</td>
<td>ल</td>
<td>-LL</td>
<td>ल ल</td>
<td>ल ल</td>
</tr>
<tr>
<td>E</td>
<td>ए</td>
<td>ए</td>
<td>-E</td>
<td>ए ए</td>
<td>ए ए</td>
</tr>
<tr>
<td>AI</td>
<td>ऐ</td>
<td>ऐ</td>
<td>-AI</td>
<td>ऐ ऐ</td>
<td>ऐ ऐ</td>
</tr>
<tr>
<td>O</td>
<td>औ</td>
<td>औ</td>
<td>-O</td>
<td>औ औ</td>
<td>औ औ</td>
</tr>
<tr>
<td>AU</td>
<td>ऑ</td>
<td>ऑ</td>
<td>-AU</td>
<td>ऑ ऑ</td>
<td>ऑ ऑ</td>
</tr>
</tbody>
</table>

Table 1: Comparison of vowel letters and signs of Siddham, Sharada, and Devanagari.
<table>
<thead>
<tr>
<th>Siddham</th>
<th>Sharada</th>
<th>Nagari</th>
<th>Siddham</th>
<th>Sharada</th>
<th>Nagari</th>
</tr>
</thead>
<tbody>
<tr>
<td>KA</td>
<td>क</td>
<td>क</td>
<td>DA</td>
<td>द</td>
<td>द</td>
</tr>
<tr>
<td>KHA</td>
<td>ख</td>
<td>ख</td>
<td>DHA</td>
<td>ध</td>
<td>ध</td>
</tr>
<tr>
<td>GA</td>
<td>ग</td>
<td>ग</td>
<td>NA</td>
<td>न</td>
<td>न</td>
</tr>
<tr>
<td>GHA</td>
<td>घ</td>
<td>घ</td>
<td>PA</td>
<td>प</td>
<td>प</td>
</tr>
<tr>
<td>NGA</td>
<td>ङ</td>
<td>ङ</td>
<td>PHA</td>
<td>फ</td>
<td>फ</td>
</tr>
<tr>
<td>CA</td>
<td>च</td>
<td>च</td>
<td>BA</td>
<td>ब</td>
<td>ब</td>
</tr>
<tr>
<td>CHA</td>
<td>छ</td>
<td>छ</td>
<td>BHA</td>
<td>भ</td>
<td>भ</td>
</tr>
<tr>
<td>JA</td>
<td>ज</td>
<td>ज</td>
<td>MA</td>
<td>म</td>
<td>म</td>
</tr>
<tr>
<td>JHA</td>
<td>झ</td>
<td>झ</td>
<td>YA</td>
<td>य</td>
<td>य</td>
</tr>
<tr>
<td>NYA</td>
<td>न</td>
<td>न</td>
<td>RA</td>
<td>र</td>
<td>र</td>
</tr>
<tr>
<td>TTA</td>
<td>ट</td>
<td>ट</td>
<td>LA</td>
<td>ल</td>
<td>ल</td>
</tr>
<tr>
<td>TTHA</td>
<td>ठ</td>
<td>ठ</td>
<td>VA</td>
<td>व</td>
<td>व</td>
</tr>
<tr>
<td>DDA</td>
<td>ड</td>
<td>ड</td>
<td>SHA</td>
<td>ढ</td>
<td>ढ</td>
</tr>
<tr>
<td>DDHA</td>
<td>ढ</td>
<td>ढ</td>
<td>SSA</td>
<td>ण</td>
<td>ण</td>
</tr>
<tr>
<td>NNA</td>
<td>ण</td>
<td>ण</td>
<td>SA</td>
<td>स</td>
<td>स</td>
</tr>
<tr>
<td>TA</td>
<td>त</td>
<td>त</td>
<td>HA</td>
<td>ह</td>
<td>ह</td>
</tr>
<tr>
<td>THA</td>
<td>थ</td>
<td>थ</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Comparison of consonant letters of Siddham, Sharada, and Devanagari.
Figure 3: Seed-syllables for the thirteen Buddhas (from Tokuzan 1974: 27).
Figure 4: The *Uṣṇīṣa Vijaya Dhāraṇī Sūtra*, known in Japanese as 佛頂尊勝陀羅尼 Buccho Sonsho Darani Kyo (Source: http://dharanipitaka.net/2008/download/). A version with vertical text is given in figure 5.
Figure 5: The Uṣṇīṣa Vijaya Dhāraṇī Sūtra written top-to-bottom, right-to-left (from Kodama 2005: 122–124).
Proposal to Encode the Siddham Script in ISO/IEC 10646

Figure 6: A manuscript containing the Prajñāpāramitā Hṛdaya Sūtra (from Müller and Nanjio 1884: Plate II). This manuscript was brought from China to Japan in 609 and was stored at the Hōryū-ji (法隆寺) temple in Ikaruga, Nara Prefecture.
Figure 7: Folio fragments from a Siddham manuscript written by Prajñātāra (Sans. प्रजातार; Chi. 般若多羅 Bōrě Duōluó) in the 9th century (from Lokesh Chandra 1965: fascicle 1, section 7). The folios are facsimiles traced by the Japanese monk Sōgen (宗源) and published using wooden-block printing in his Ashara-jō (1837).
Figure 8: Siddham inscriptions (from Nakamura 1977: Plate 21)
Figure 9: Siddham inscriptions (from Nakamura 1977: Plate 30)
Figure 10: The Siddham varṇapāṭha written by Kūkai (from van Gulik 1980: Plate XI)
Figure 11: A maṇḍala inscribed by Saichō (from van Gulik 1980: Plate X)
### VOWELS (MĀTĀ)

<table>
<thead>
<tr>
<th></th>
<th>formal</th>
<th>brush</th>
<th>stroke order</th>
<th>variations</th>
<th>in combination</th>
<th>Devanāgarī</th>
<th>roman letter</th>
<th>esoteric meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ँ</td>
<td>मँ</td>
<td></td>
<td></td>
<td>अ</td>
<td>अ</td>
<td>a</td>
<td>anupāda unborn</td>
</tr>
<tr>
<td>2</td>
<td>ं</td>
<td>मं</td>
<td></td>
<td></td>
<td>आ</td>
<td>ā</td>
<td>a</td>
<td>akāśa space</td>
</tr>
<tr>
<td>3</td>
<td>ः</td>
<td>मः</td>
<td></td>
<td></td>
<td>इ</td>
<td>i</td>
<td>i</td>
<td>indriya sense organ</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ई</td>
<td>ī</td>
<td>i</td>
<td>īti calamity</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>उ</td>
<td>u</td>
<td>u</td>
<td>upamā simile</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ऊ</td>
<td>ū</td>
<td>u</td>
<td>ūna incomplete</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ए</td>
<td>e</td>
<td>e</td>
<td>esāpa seeking</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ऐ</td>
<td>ai</td>
<td>ai</td>
<td>aśvarya sovereignty</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ओ</td>
<td>o</td>
<td>o</td>
<td>ogha flood</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ऐ</td>
<td>au</td>
<td>au</td>
<td>aupapādūka self-produced</td>
</tr>
</tbody>
</table>
### Figure 13: Description of Siddham signs Anusvara and Visarga, the vocalic letters, and the consonants Ka, Kha (from Stevens 1981: 35).

<table>
<thead>
<tr>
<th>#</th>
<th>Formal</th>
<th>Brush</th>
<th>Stroke Order</th>
<th>Variations</th>
<th>In Combination</th>
<th>Devanagari</th>
<th>Roman Letter</th>
<th>Esoteric Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>ड</td>
<td>ड़</td>
<td>स्ख</td>
<td>स्खँ</td>
<td>सँ</td>
<td>a m</td>
<td>anta limit</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>ढ</td>
<td>ढ़</td>
<td>स्ध</td>
<td>स्धँ</td>
<td>सँ</td>
<td>a h</td>
<td>astapagama setting</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>ठ</td>
<td>ठ़</td>
<td>स्ठ</td>
<td>स्ठँ</td>
<td>सँ</td>
<td>ठि</td>
<td>tāḍhī supernatural power</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>ड़</td>
<td>ड़़</td>
<td>स्ड़</td>
<td>स्ड़ँ</td>
<td>सँ</td>
<td>ड़ि</td>
<td>analogy</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>ल</td>
<td>ल़</td>
<td>स्ल़</td>
<td>स्ल़ँ</td>
<td>सँ</td>
<td>ल़ि</td>
<td>dye</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>ल़</td>
<td>ल़़</td>
<td>स्ल़</td>
<td>स्ल़ँ</td>
<td>सँ</td>
<td>ल़़ि</td>
<td>submerge</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>क</td>
<td>क़</td>
<td>स्क़</td>
<td>स्कँ</td>
<td>सँ</td>
<td>क</td>
<td>karma action</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>क्ष</td>
<td>क्ष़</td>
<td>स्क्ष़</td>
<td>स्क्ष़ँ</td>
<td>सँ</td>
<td>क्ष</td>
<td>kha sky</td>
<td></td>
</tr>
<tr>
<td>formal</td>
<td>brush</td>
<td>stroke order</td>
<td>variations</td>
<td>in combination</td>
<td>Devanāgarī</td>
<td>roman letter</td>
<td>esoteric meaning</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>--------------</td>
<td>------------</td>
<td>----------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>ओ</td>
<td></td>
<td></td>
<td>र</td>
<td>ग</td>
<td>ga</td>
<td>gati, going</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>क</td>
<td></td>
<td></td>
<td>घ</td>
<td>घ</td>
<td>gha</td>
<td>ghan, dense</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>क</td>
<td></td>
<td></td>
<td>ञ</td>
<td>ञ</td>
<td>na</td>
<td>aṅga, part</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>क</td>
<td></td>
<td></td>
<td>च</td>
<td>च</td>
<td>ca</td>
<td>cyuti, transition</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>क</td>
<td></td>
<td></td>
<td>छ</td>
<td>छ</td>
<td>cha</td>
<td>chaiva, shadow</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>क</td>
<td></td>
<td></td>
<td>ज</td>
<td>ज</td>
<td>ja</td>
<td>jāti, birth</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>क</td>
<td></td>
<td></td>
<td>झ</td>
<td>झ</td>
<td>jha</td>
<td>jhaṣabala, warring enemies</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>क</td>
<td></td>
<td></td>
<td>ञ</td>
<td>ञ</td>
<td>na</td>
<td>jñāna, knowledge</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>क</td>
<td></td>
<td></td>
<td>त</td>
<td>त</td>
<td>ta</td>
<td>pāṇka, pride</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>क</td>
<td></td>
<td></td>
<td>ठ</td>
<td>ठ</td>
<td>tha</td>
<td>viṭhapana, flourish</td>
<td></td>
</tr>
</tbody>
</table>

Figure 14: Description of Siddham consonants GA .. THA (from Stevens 1981: 36).
<table>
<thead>
<tr>
<th>formal</th>
<th>brush</th>
<th>stroke order</th>
<th>variations</th>
<th>combination</th>
<th>Devanagari</th>
<th>roman letter</th>
<th>esoteric meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ड</td>
<td>da</td>
<td>damara tumult</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ढ</td>
<td>ḍha</td>
<td>disappear</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ण</td>
<td>ṇa</td>
<td>battle</td>
</tr>
<tr>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>त</td>
<td>ta</td>
<td>tathātā suchness</td>
</tr>
<tr>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>थ</td>
<td>tha</td>
<td>dwelling</td>
</tr>
<tr>
<td>34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>द</td>
<td>da</td>
<td>dāna generosity</td>
</tr>
<tr>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ध</td>
<td>dha</td>
<td>dharma realm</td>
</tr>
<tr>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>न</td>
<td>na</td>
<td>nāman name</td>
</tr>
<tr>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>प</td>
<td>pa</td>
<td>paramāṇtha ultimate meaning</td>
</tr>
<tr>
<td>38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>फ</td>
<td>pha</td>
<td>phena foam</td>
</tr>
</tbody>
</table>

Figure 15: Description of Siddham consonants dda .. pha (from Stevens 1981: 37).
<table>
<thead>
<tr>
<th>formal</th>
<th>brush</th>
<th>stroke order</th>
<th>variations</th>
<th>in combination</th>
<th>Devanagari</th>
<th>roman letter</th>
<th>esoteric meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
<td>ब</td>
<td>बा</td>
<td>ब</td>
<td>ba</td>
<td>बा</td>
<td>ba</td>
<td>bandhana binding</td>
</tr>
<tr>
<td>40</td>
<td>भ</td>
<td>भा</td>
<td>भ</td>
<td>bha</td>
<td>भा</td>
<td>bha</td>
<td>existence</td>
</tr>
<tr>
<td>41</td>
<td>म क</td>
<td>मा</td>
<td>म</td>
<td>ma</td>
<td>मा</td>
<td>ma</td>
<td>mama my</td>
</tr>
<tr>
<td>42</td>
<td>य क</td>
<td>या</td>
<td>य</td>
<td>ya</td>
<td>या</td>
<td>ya</td>
<td>vehicle</td>
</tr>
<tr>
<td>43</td>
<td>र ड</td>
<td>रा</td>
<td>र</td>
<td>ra</td>
<td>रा</td>
<td>ra</td>
<td>rajas passion</td>
</tr>
<tr>
<td>44</td>
<td>ल क</td>
<td>ला</td>
<td>ल</td>
<td>la</td>
<td>ला</td>
<td>la</td>
<td>lakṣaṇa mark</td>
</tr>
<tr>
<td>45</td>
<td>व क</td>
<td>वा</td>
<td>व</td>
<td>va</td>
<td>वा</td>
<td>va</td>
<td>vāc speech</td>
</tr>
<tr>
<td>46</td>
<td>स क</td>
<td>सा</td>
<td>श</td>
<td>sa</td>
<td>सा</td>
<td>sa</td>
<td>śānti peace</td>
</tr>
<tr>
<td>47</td>
<td>ष क</td>
<td>षा</td>
<td>ष</td>
<td>sa</td>
<td>षा</td>
<td>sa</td>
<td>śādāyatana six senses</td>
</tr>
<tr>
<td>48</td>
<td>स क</td>
<td>सा</td>
<td>स</td>
<td>sa</td>
<td>सा</td>
<td>sa</td>
<td>satya truth</td>
</tr>
</tbody>
</table>

Figure 16: Description of Siddham consonants बा .. सा (from Stevens 1981: 38).
Figure 17: Description of Siddham consonant letter \( \text{ha} \) and the conjuncts \( \text{kṣa} \) and \( \text{llam} \) (from Stevens 1981: 39).

Figure 18: A set of ‘Bonji’ digits (Source: http://www.mandalar.com/DisplayJ/Bonji/index6_E.html). These closely resemble Devanagari forms.
<table>
<thead>
<tr>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
<th>Column 5</th>
<th>Column 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
| Figure 19: Chart of Siddham (from Kodama 1991: 65–66). Note the variant form of ꔲ.

Figure 19: Chart of Siddham (from Kodama 1991: 65–66). Note the variant form of ꔲ.
Proposal to Encode the Siddham Script in ISO/IEC 10646

Anshuman Pandey

Figure 20: Chart of Siddham (from Kodama 1991: 67–68).
Figure 21: Siddham letters used in Korea (from Lim 2005). Courtesy of Lee Collins.
Figure 22: Siddham letters used in Korea (from Lim 2005). Courtesy of Lee Collins.
Figure 23: Siddham letters used in Korea (from Lim 2005). Courtesy of Lee Collins.
Proposal to Encode the Siddham Script in ISO/IEC 10646

Anshuman Pandey

Figure 24: Siddham letters used in Korea (from Lim 2005). Courtesy of Lee Collins.
Figure 25: Siddham used in a Korean version of the Nīlakaṇṭha Dhāranī (नीलकण्ठ धारनी; Chi. 大悲咒 Dàbēi Zhòu) of Avalokiteśvara (from Lokesh Chandra 1988: 164).
Figure 26: Siddham letters used in Korea (from Stevens 1981: 8)
Figure 27: Chart showing the variety of Siddham used in the Hōryū-ji palm-leaf manuscripts (top portion of figure) (from Ojhā 1971: Plate 19).
This type of consonants and vowels are also available in various inscriptions of Nepal.

Figure 28: Chart showing the variety of Siddham used in Nepal and Central Asia (from Śākyavāṃśa 1985: 38–39). This form of Siddham is apparently that used on the Hōryū-ji palm-leaf manuscripts.
Figure 29: Comparison of Siddham, Tibetan, Ranjana, and Chinese scripts (Source: http://www.siddham-sanskrit.com/s-sanskrit2/learn-siddham/sid-ran-tib-han.GIF).
Figure 30: Description of the basic vowels, anusvara, and visarga in the Shittan Sanmitsushō (from Jōgon 1682: 722). Note the several glyphic variants given for each vowel.
Figure 31: Description of vowel signs and the vocalic letters in the *Shittan Sanmisushō* (from Jōgon 1682: 722).
Punctuation marks are as follows:

-  is written at the beginning of a text
-  ditto mark
-  comma (formal)
-  comma (informal)
-  period (formal)
-  period (informal)

At the end of a text these marks are often used:

Figure 32: Punctuation and ornaments used in Siddham (from Stevens 1981: 41).

Figure 33: Punctuation and ornaments used in Siddham (from Tokuzan 1974: 227). Many of these are identical to those shown by Stevens (figure 32) and Kodama (figure 35). However, note here the vertical variant \( \text{\textcircled{}} \) of \( \text{\textcircled{}} \) SIDDHAM.
Figure 34: Description of an invocation sign, punctuation marks, section marks, a variant form of virama, a deletion mark, and repetition marks in the Shittan Sanmitsushō (from Jōgon 1682: 724). Note the separate listings for Ⅷ SECTION MARK EIGHT and Ⅸ SECTION MARK ELEVEN.
Figure 35: Punctuation and ornaments used in Siddham (from Kodama 1991: 222). Column (1) forms of virama; (2) the siddham sign; (3) the ‘informal’ danda; (4) the ‘informal’ double danda; (5) the three repetition marks; (6) deletion marks; (7) ornaments used at the end of a text section; (8) a sign used at the end of text.
Figure 36: Chinese Siddham showing usage of 2 REPETITION MARK ONE (source unknown). Note the serialized use of 2 after the syllable hūṃ, which indicates that the word is to be read thrice.
Figure 37: Excerpt from the *Uṣṇīṣa Vijaya Dhāraṇī Sūtra* in figure 5 showing usage of \*\*\*REPETITION MARK TWO\*\*\* in vertical text (from Kodama 2005: 123).
Figure 38: Use of ꠊ WORD SEPARATOR ONE (from Lokesh Chandra 1965: fascicle 1, section 12). Note also the use of the invocation sign ꠮ SIDDHAM at the beginning of the text.

Figure 39: Use of a middle dot as a word separator (from Lokesh Chandra 1972: 410). Note also the use of ꠊ SECTION MARK TWELVE and ꠊ SECTION MARK THIRTEEN.
Figure 40: Text showing use of 𑗂 DANDA, 𑗃 DOUBLE DANDA, and 𑗅 WORD SEPARATOR TWO (from Lokesh Chandra 1972: 15). Note the writing of DANDA and DOUBLE DANDA beneath WORD SEPARATOR TWO.

Figure 41: Text showing use of 𑗅 WORLD SEPARATOR TWO, 𑗂 DANDA, and 𑗃 DOUBLE DANDA (from Lokesh Chandra 1972: 33). Here WORLD SEPARATOR TWO is written independently.
The mantra of Pāṇḍaravāsī: namaha samanta-buddhānāṁ paṁ. 90

The mantra of Hayagrīva: namaha samanta-buddhānāṁ harṁ. 91

The mantra of Yasodharā: namaha samanta-buddhānāṁ yaṁ. 92

The mantra of Ratnapāṇi: namaha samanta-buddhānāṁ saṁ. 93

The mantra of Jālinīprabha: namaha samanta-buddhānāṁ jauṁ. 94

The mantra of Śākyamuni: namaha samanta-buddhānāṁ bhauṁ. 95

The mantra of the three Bodhisattvas: namaha samanta-buddhānāṁ haṁ truṁ. 96

The mantra of Śītāpatroṣṭa: namaha samanta-buddhānāṁ tuṁ. 97

The mantra of Jayoṣṭha: namaha samanta-buddhānāṁ saṁ. 98

The mantra of Vijayoṣṭha: namaha samanta-buddhānāṁ st. 99

Figure 42: Text showing use of । word separator two and ।। double danda (from Yamamoto 1990: 61, 62).
Figure 43: Chart showing the use of ꝏ NUKTA for writing sounds not natively represented by traditional Siddham letters (Source: http://www.mandalar.com/DisplayJ/Bonji/index2.html).
Figure 44: Text showing the use of ꝫ NUKTA with Ꝫ JA in order to represent /z/ in the Siddham for the Japanese word 刺青 irezumi “tattoo” (Source: http://www.mandalar.com/BonjiDigitalDictionarySAMPLE/member/_Tattoo/00Tattoo.html).
Figure 4.5: Grid showing traditional Siddham character set (adapted from Stevens 1981: 40–41). This appears to be based upon the analysis of Siddham conjuncts by Zhī-guǎng (see Section 3.9).
Figure 46: Siddham ligatures (from van Gulik 1980: Plate VIIa).

Figure 47: Siddham ligatures (from van Gulik 1980: Plate VIIb).
Plate VIIc. Siddham ligatures

Plate VII d. Siddham ligatures

Figure 48: Siddham ligatures (from van Gulik 1980: Plate VIIc).

Figure 49: Siddham ligatures (from van Gulik 1980: Plate VII d).
Plate VII e. Siddham ligatures

Plate VII f. Siddham ligatures

Figure 50: Siddham ligatures (from van Gulik 1980: Plate VIIe).

Figure 51: Siddham ligatures (from van Gulik 1980: Plate VII f).
Figure 52: Siddham ligatures (from van Gulik 1980: Plate VIIg).

Figure 53: Siddham ligatures (from van Gulik 1980: Plate VIIh).
Plate VII.i. Siddham ligatures

Plate VII.j. Siddham ligatures

Figure 54: Siddham ligatures (from van Gulik 1980: Plate VIIi).

Figure 55: Siddham ligatures (from van Gulik 1980: Plate VIIj).
Figure 56: Siddham ligatures (from van Gulik 1980: Plate VIIk).

Figure 57: Siddham ligatures (from van Gulik 1980: Plate VIIl).
Figure 58: Siddham ligatures (from van Gulik 1980: Plate VIIm).

Figure 59: Siddham ligatures (from van Gulik 1980: Plate VIIIn).
Figure 60: Siddham ligatures (from van Gulik 1980: Plate VIIo).

Figure 61: Siddham ligatures (from van Gulik 1980: Plate VIlp).
Figure 62: Chart showing ‘headless’ consonants combined with virama (from Jōgon 1682: 779). These are pedagogical forms of regular consonants letters (see Section 4.4).
Figure 63: Chart showing Siddham dependent vowel signs and their combinations with क्य (from Shinpan: 644). Note the presence of the vowel sign  for आ, which is derived from the top-stroke of the letter क्य. See Section 4.4 for more details.

Figure 64: The  "myō-ten" shown as the dependent form of ऋ vowel letter आ (from Tokuzan 1974: 226).
Figure 65: Elemental Siddham strokes (from Bonji Kichō Shiryō Kankōkai 1980: 114).
Figure 66: Stroke sequence for writing घ using a soft brush (from Stevens 1981: 44). The initial stroke ई myō-ten is highlighted. See Section 4.4 for details on Siddham stroke characters.

Figure 67: The position of ई myō-ten in the seed-syllables for the five elements आ, ए, ऐ, क्ष, छ (from Tokuzan 1974: 180).
Figure 68: A conceptualization of the syllable वा (va) (from Tokuzan 1976: 43). The elements from top are च, ज, शा, स्त, ट्या, conceptualized as a stūpa (from Tokuzan 1976: 28).

Figure 69: The seed-syllables for the five elements च, ज, शा, स्त, ट्या, conceptualized as a stūpa (from Tokuzan 1976: 42). The elements from top are च, ज, शा, स्त, ट्या.
Figure 70: Different forms of hāṃmām, the bija of Fudō Myōō (Acala) (from Tokuzan 1974: 227).

Figure 71: The seed syllables of Fudō Myōō and his attendants,制吒迦童子 Seitaika-dōji (Ceṭaka) and 矜羯羅童子 Kongara-dōji (Kiṃkara) (from Kodama 2005: 97).
Figure 72: Two variations of the seed syllable dhīhīṃ (from Kodama 2005: 99).
**A. Administrative**

1. **Title:** Proposal to Encode the Siddham Script 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:** 2012-07-23
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): Yes
   - Proposed name of script: Siddham
   - b. The proposal is for addition of character(s) to an existing block:
   - Name of the existing block:

2. **Number of characters in proposal:** 85

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
   - X

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 and Lee Collins (Apple)
   - b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
     - Lee Collins (Apple)

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

   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

   If YES, with whom?  
   Shingon Buddhist community, Indic script experts

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?  
   Yes

   Reference:

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

   Reference:

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

   Reference:

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?  
   N/A

   If YES, is a rationale provided?  
   If YES, reference:

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

8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?  
   No

   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?  
   No

   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, or could be confused with, an existing character?  
    Yes

    If YES, is a rationale for its inclusion provided?  
    Yes

    If YES, reference:  
    Virama, etc.; see text of proposal

11. Does the proposal include use of combining characters and/or use of composite sequences?  
    Yes

    If YES, reference:  
    Dependent vowel signs, etc.; see text of proposal

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

    If YES, reference:

12. Does the proposal contain characters with any special properties such as control function or similar semantics?  
    Yes

    If YES, describe in detail (include attachment if necessary)  
    Virama, etc.; see text of proposal

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

    If YES, are the equivalent corresponding unified ideographic characters identified?  
    If YES, reference: