

**Title:** Proposal to Encode a Set of Digits for Siddham  
**Source:** Scripts Encoding Initiative (SEI)  
**Author:** Anshuman Pandey (pandey@umich.edu)  
**Status:** Liaison Contribution  
**Action:** For consideration by UTC and WG2  
**Date:** 2013-06-15

## 1 Introduction

This is a proposal to encode the following ten characters in the Siddham block of the Universal Character Set (ISO/IEC 10646):

GLYPH	CODE	CHARACTER NAME	GLYPH	CODE	CHARACTER NAME
ୠ	115E0	SIDDHAM DIGIT ZERO	ୡ	115E5	SIDDHAM DIGIT FIVE
ୢ	115E1	SIDDHAM DIGIT ONE	ୣ	115E6	SIDDHAM DIGIT SIX
୤	115E2	SIDDHAM DIGIT TWO	୦	115E7	SIDDHAM DIGIT SEVEN
୥	115E3	SIDDHAM DIGIT THREE	୧	115E8	SIDDHAM DIGIT EIGHT
୦	115E4	SIDDHAM DIGIT FOUR	ୡ	115E9	SIDDHAM DIGIT NINE

The Siddham script was proposed for encoding by the present author (see N4294 L2/12-234R) and accepted for inclusion in the UCS at the meeting of the Unicode Technical Committee in August 2012. These characters were identified in original proposal for Siddham, but were not proposed for inclusion because additional information regarding their usage was required.

## 2 Background

Digits are not attested in historical Siddham records. However, there are Siddham characters whose form and function indicate their origins from digits, such as ୡ SIDDHAM REPETITION MARK-1 and ୣ SIDDHAM REPETITION MARK-3. These marks are written after a word that is to be repeated during recitation. Based upon the shape and usage of these characters, it is likely that they are derived from an Indic digit ‘2’ (compare Devanagari २). This is further evident through an analysis of variants of ୡ SIDDHAM REPETITION MARK-1, which have the forms ୠ and ୢ. A common practice in manuscript traditions of northern India is to indicate the repetition of words using digits. The choice of digit indicates the number of repetitions, eg. ‘2’ for ‘twice’, ‘3’ for ‘thrice’. It is plausible that in the Siddham tradition the linkage between the repetition marks and their numerical origin was forgotten and the digit for ‘2’ was preserved as a generic mark of repetition. This explains the serialized usage of REPETITION MARK-1 in Siddham texts for indicating a triple reading, compare “ୠ ୡ ୡ” with Devanagari “३ ३” for “*hūṃ hūṃ hūṃ*”).

In order to compensate for the absence of digits in Siddham, contemporary users have designed digits that are based upon the graphical model of northern Indic digits, namely Devanagari. These modern innovations are called “Bonji” digits, *bonji* (梵字) being a popular name for the Siddham script in Japan.

### 3 Character Data

#### 3.1 Character Properties

Properties for Siddham digits in the Unicode Character Database format are:

```
115E0;SIDDHAM DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
115E1;SIDDHAM DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
115E2;SIDDHAM DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
115E3;SIDDHAM DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
115E4;SIDDHAM DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
115E5;SIDDHAM DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
115E6;SIDDHAM DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
115E7;SIDDHAM DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
115E8;SIDDHAM DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
115E9;SIDDHAM DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
```

#### 3.2 Linebreaking

Linebreaking properties of Siddham digits given in the data format of `LineBreak.txt`:

```
115E0..115E9; NU # SIDDHAM DIGIT ZERO .. SIDDHAM DIGIT NINE
```

### 4 References

Pandey, Anshuman. 2012. “Proposal to Encode the Siddham Script in ISO/IEC 10646”. ISO/IEC JTC1/SC2/WG2 N4294 L2/12-234R. August 1, 2012. <http://std.dkuug.dk/jtc1/sc2/wg2/docs/n4294.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.

Numeral		English sound							
Sanskrit-Bonji									
0	shuunya	1	eeka	2	dvi	3	tri	4	chatur
०		१		२		३		४	
s'u:nyā		e+ka		dvi		tri		c'a+tu+r	
𑖀𑖁		𑖂𑖃		𑖄𑖅		𑖆𑖇		𑖈𑖉𑖊	
5	pancha	6	shash	7	sapta	8	ashta	9	nava
𑖋		𑖌		𑖍		𑖎		𑖏	
pam,c'a		s,a+s̄		sa+pta		a+s,t,a		na+va	
𑖐𑖑		𑖒𑖓		𑖔𑖕		𑖖𑖗		𑖘𑖙	
Bonji-Character Pronunciation Diagram									

Figure 1: A set of 'Bonji' digits with Sanskrit names for the digits written in Siddham (Source: [http://www.mandalar.com/DisplayJ/Bonji/index6\\_E.html](http://www.mandalar.com/DisplayJ/Bonji/index6_E.html)).