**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 2/WG 2**

**Universal Multiple-Octet Coded Character Set
(UCS)**

ISO/IEC JTC 1/SC 2/WG 2 *N4407*
2013-04-20

Title:	**Proposal to Encode Variants for Siddham Script**
Source:	**Taichi KAWABATA, Toshiya SUZUKI, Kiyonori NAGASAKI and Masahiro SHIMODA**
References:	**See "References" section in the document**
Status:	**Experts' Contribution**
Action:	**Request for review by ISO/IEC JTC 1/SC 2/WG 2**

Proposal to Encode Variants for Siddham Script

Taichi KAWABATA, Toshiya SUZUKI, Kiyonori NAGASAKI and Masahiro SHIMODA

Abstract

This document proposes a set of variation sequences for Siddham script. Encoding of Siddham script was proposed by Anshuman Pandey as WG2 Document N4294 ("*Proposal to Encode the Siddham Script*"), and these characters are planned to be published as ISO/IEC 10646:2012 Amendment 2 (see M60.3 of N4354), which is currently at a stage of PDAM2. PDAM2 unifies known Siddham variants into a single character, and it is sane and useful, as described in the proposal of N4294. On the other hand, in Japan, several Siddham variant glyphs are sometimes distinctively used to denote different religious icons. In this document, authors would like to propose a use of v*ariation selector mechanism* to distinguish them in such case.

Background

Siddham script is one of Brahmi scripts that was born in 6th or 7th century of India. However, its modern use is limited to Japanese Buddhism and related research activities.

In Buddhism, Siddham scripts are used not only as phonetic characters, but also as symbols for some form of bodhisattvas. By this mean, Siddham characters can be regarded as ideograms. Siddham script has several kinds of variants that are phonetically equivalent. However, some variants may denote different meanings when they are used as ideograms. For example, there are at least three phonetically

equivalent variants for short vowel "I" (⟨glyph⟩, ⟨glyph⟩, ⟨glyph⟩), however, they may be treated distinctively when used

as ideograms. For example, as shown in Figure 7 below, ⟨glyph⟩ denotes *Shoryu (*諸龍*)*, while ⟨glyph⟩ denotes *Zenrin (*禅林*)*, that are different manifestation of Ksiti Garbha Bodhisattva.

Following tables summarizes Siddhams as ideograms shown in the following figures. It should be noted that subjects of symbolization by these ideograms may vary among the documents or Buddhist sects, but subjects are consistent within a single document, and the shape of these variants are also stable and consistent among the documents and sects.

Glyphs	Character	Name of Kstigrabhas or others in each figure with the symbol.
⟨glyph⟩	I	金剛宝（Figure 5）, 鶏亀（Figure 6）, 諸龍 （Figure 7）, 金剛願·光味（Figure 8）, 金剛源·諸龍（Figure 9）, 金剛源（Figure 10）
⟨glyph⟩	I	放光王 (Figure 5）, 宝性 （Figure 6）, 禅林 （Figure 7）, 放光·破勝（Figure 8）, 放光·禅林（Figure 9）, 放光王（Figure 10）

	I	讚龍 (Figure 4) 讚龍（Figure 8）
	Ī	黒 (Figure 4) 地持地蔵（Figure 6）伏勝（Figure 7）金剛宝（Figure 9）
	Ī	弁尼 (Figure 4) 法印地蔵（Fig 6）
	U	SIDDHAM LETTER U form 1 烏波難陀龍王 (Figure 2)
	U	SIDDHAM LETTER U form 2 大自在天妃 (Figure 2)
	U	降三世明王 (Figure 1, 11, 12, 13).
	U	軍荼利明王 (Figure 1, 11, 12, 13)
	Ū	Differently treated symbol at Mandala (Figure 14)
	Ū	Differently treated symbol at Mandala (Figure 14)

Proposal

Authors would like to propose the following variation selectors for variants. We have identified that these variants sometimes convey semantic differences, which should not be expressed by the fonts.

Glyphs	Variation Sequences	Description
	<11582 FE00>	SIDDHAM LETTER I form 1
	<11582 FE01>	SIDDHAM LETTER I form 2
	<11582 FE02>	SIDDHAM LETTER I form 3
	<11583 FE00>	SIDDHAM LETTER II form 1
	<11583 FE01>	SIDDHAM LETTER II form 2
	<11584 FE00>	SIDDHAM LETTER U form 1
	<11584 FE01>	SIDDHAM LETTER U form 2

	<115B2 FE00>	SIDDHAM VOWEL SIGN U form 1
	<115B2 FE01>	SIDDHAM VOWEL SIGN U form 2
	<115B3 FE00>	SIDDHAM VOWEL SIGN UU form 1
	<115B3 FE01>	SIDDHAM VOWEL SIGN UU form 2

Goals and Non-Goals of This Proposal

Variants in this proposal are intended be used when the one needs to distinguish the semantic differences of the same character as an ideogram, as described in the above text. This proposal does not intend to provide a way to distinguish the regional, religious or historical differences of variants.

Additional Notes

The proposal includes some (actually four) variation sequences to specify the variation glyph for combining characters (Siddham vowel signs). Such variation sequences appear to be disallowed in the current UCS standard, as 16.5 says:

> Variation selectors are a specific class of combining characters immediately following a non decomposable base character

However, this restriction is technically too strong, and there should be no technical problem if variation selectors are allowed to be specified to combining characters whose combining class is equal to zero. This proposal assumes that UCS standards to be changed so that combining characters with combining class equal to zero can have the variation selectors. To be concrete, changing the above statement to something like the following statement, for example, would be appropriate.

> Variation selectors are a specific class of combining characters immediately following a non decomposable character whose combining class is not equal to zero.

Acknowledgements

Authors are thankful to Anshuman Pandey and SBII for their eager work for the original proposal of Siddham script and would like to thank Prof. Giryu Kodama (Shuchiin University) for his comprehensive advice and providing the valuable evidences and The SAT Daizōkyō Text Database Committee (led by Prof. Masahiro Shimoda, University of Tokyo) especially Prof. Koju Motoyama (Taisho University) and Dr. Kenichi Kurahisni (Taisho University) regarding providing of their expertise.

References

Anshuman Pandey (2012), "Proposal to Encode the Siddham Script in ISO/IEC 10646", JTC 1/SC 2/WG 2 N4294

Bibliography of Figures

Tokuyama, Kijun (徳山暉純) , Illustration and Explanation of Siddham Script (図説 梵字), Published by Mokujisha (木耳社), 1974
Somekawa, Eisuke (染川英輔), Illustrations of Mandala (曼荼羅図典), 1993
Kodama, Giryu (児玉義隆) (1991), Comanion for Siddham Script (梵字必携)
Kodama, Giryu (児玉義隆) (2002), Daiho-Rinkaku (大法輪閣)
Tendai-Sect (天台宗) Comanion for Buddhism, Revised edition (改訂仏事必携版)
『大正新脩大藏經』 [Taisho-Shinshu-Daizokyo], 1924-1934

『大正新脩大藏經 図像部』[Taisho-Shinshu-Daizokyo Zuzoubu], 1924-1934
Classical books from the collection of Prof. Giryu Kodama
『諸尊種字 [Shoson-Shuji] 』
『諸真言要集[Shoshingon-youshu] 』巻一 (Vol.1)

不動曼荼羅・中央尊

Figure 1. Two variants of SIDDHAM VOWEL SIGN U are semantically different.
(from ［Tokuyama：1974］, p. 34)

In the above figure, two variants of vowel sign behave as different ideograms, as shown in the following figure.

Figure 2. Two variants of SIDDHAM LETTER U are semantically different.
(Excerpt from 染川英輔(Somekawa, Eisuke)『曼荼羅図典』(Mandara-zuten), 1993)

Figure 3. Two variants of SIDDHAM LETTER II have different semantics.
(from 『大正新脩大藏經』(Taisho-Shinshu-Daizokyo), 1924-1934, vol. 84, p. 407)
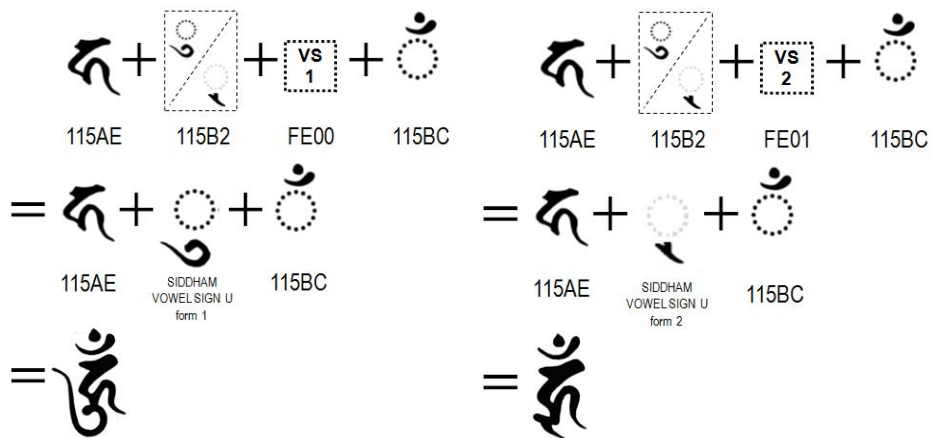
Figure 4. Three variants of SIDDHAM LETTER I and two variants of SIDDHAM LETTER II have respectively different semantics.
(from [Kodama:1991], p. 233.)

Figure 5. Two variants of SIDDHAM LETTER II have different semantics.
(from 『諸尊種字(Shoson-Shuji)』, A collection of Prof. Giryu Kodama)
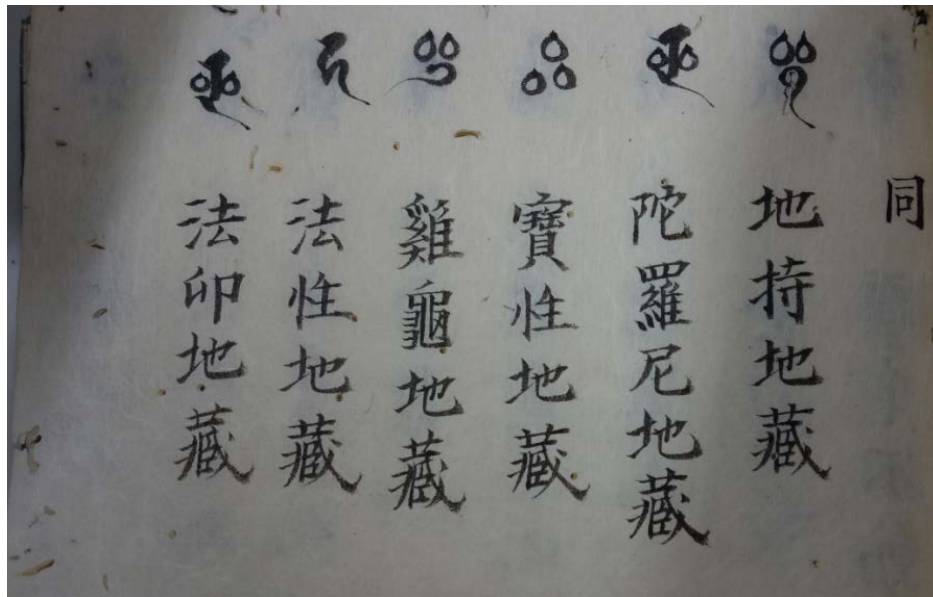
Figure 6. Two variants of SIDDHAM LETTER I and two variants of SIDDHAM LETTER II have
respectively different semantics.
(from 『諸尊種字(Shoson-Shuji)』, A collection of Prof. Giryu Kodama)
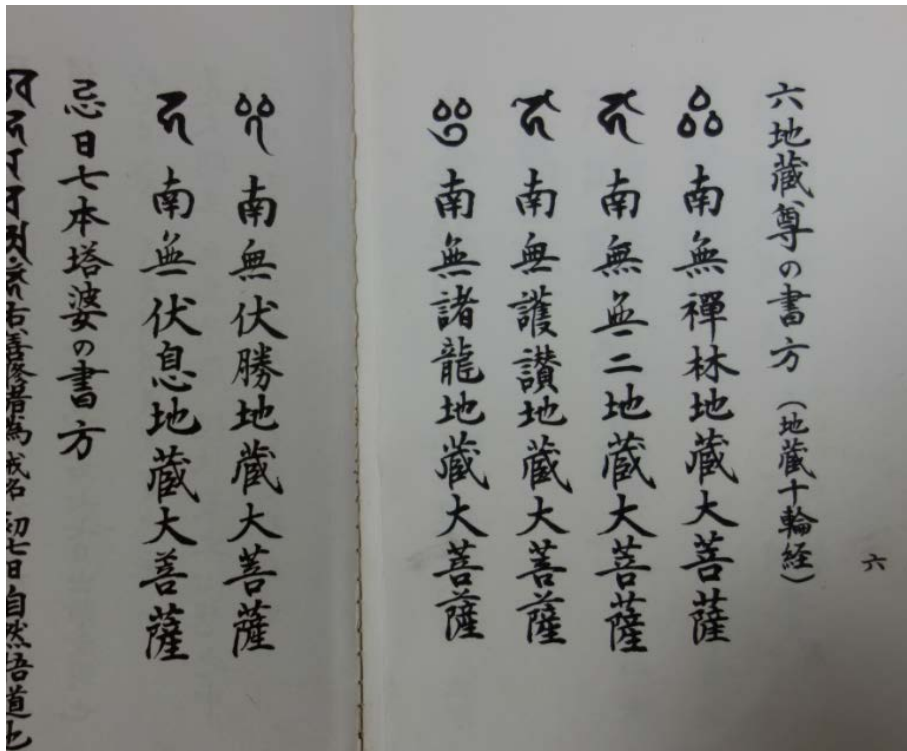
Figure 7. Two variants of SIDDHAM LETTER I have different semantics.
(from a textbook of the Chizan school)

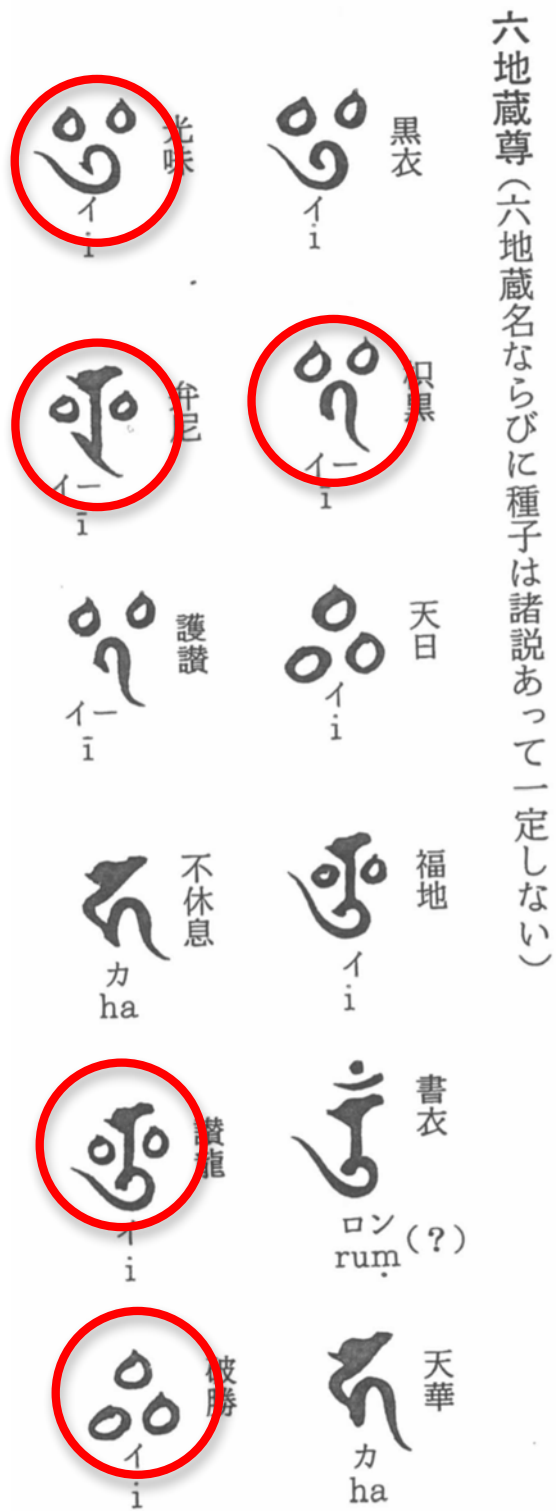Figure 8. Three variants of SIDDHAM LETTER I and two variants of SIDDHAM LETTER II have
respectively different semantics.
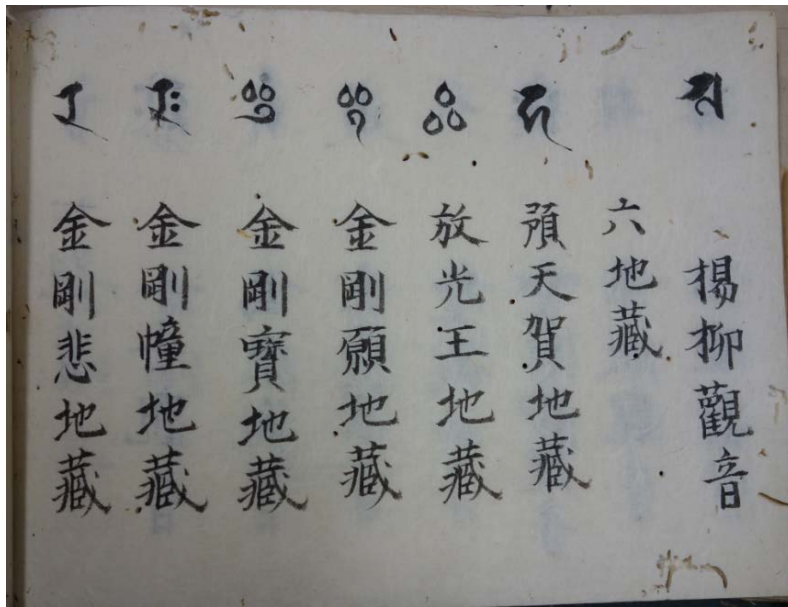(from 『諸真言要集(Shoshingon-youshu)』, A collection of Prof. Giryu Kodama)
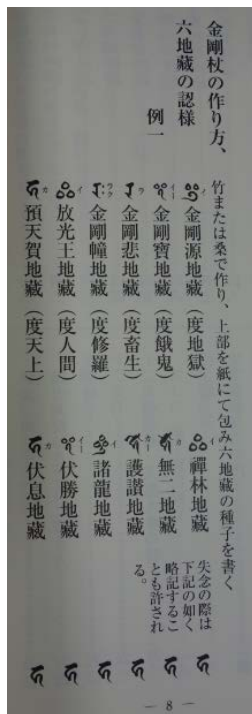
Figure 9. Two variants of SIDDHAM LETTER I have different semantics.
(from 天台宗(Tendai-shu)『改訂仏事必携』(Kaitei-butsuji-hikkei))
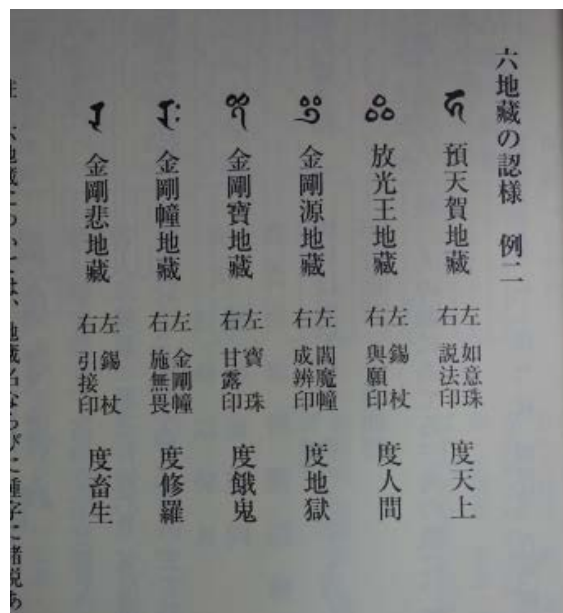
Figure 10. Two variants of SIDDHAM LETTER I have different semantics.
(from 天台宗(Tendai-shu)『改訂仏事必携』(Kaitei-butsuji-hikkei))
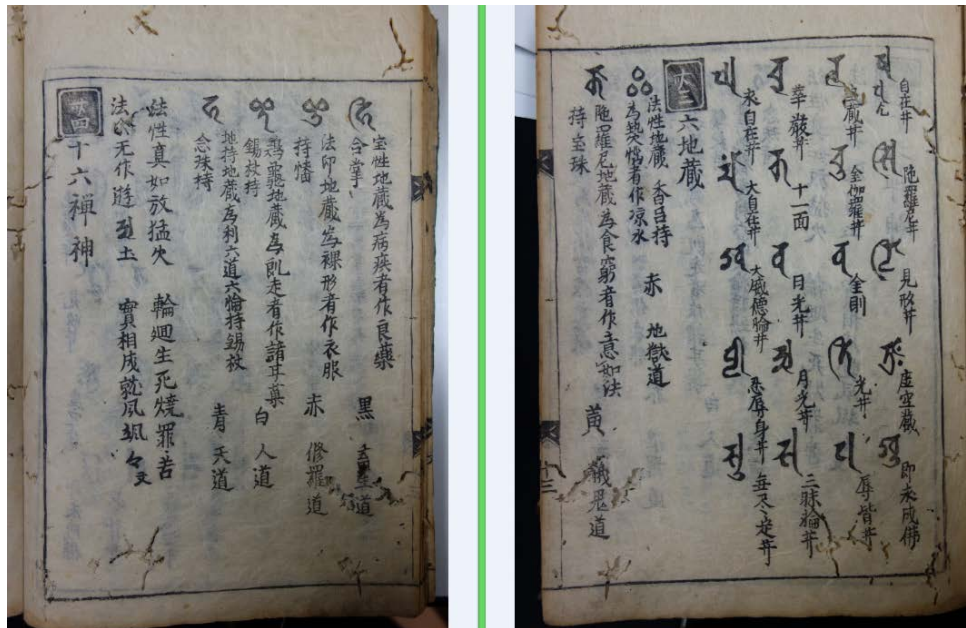
Figure 11. Two variants of SIDDHAM LETTER I have different semantics.
(from 『諸真言要集 [Shoshingon-youshu] 』巻一 (Vol.1))

Figure 12. SIDDHAM VOWEL SIGN U are semantically different.
(from 『大正新脩大藏經 図像部』(Taisho-Shinshu-Daizokyo Zuzoubu), 1933, vol. 1, p. 484.)
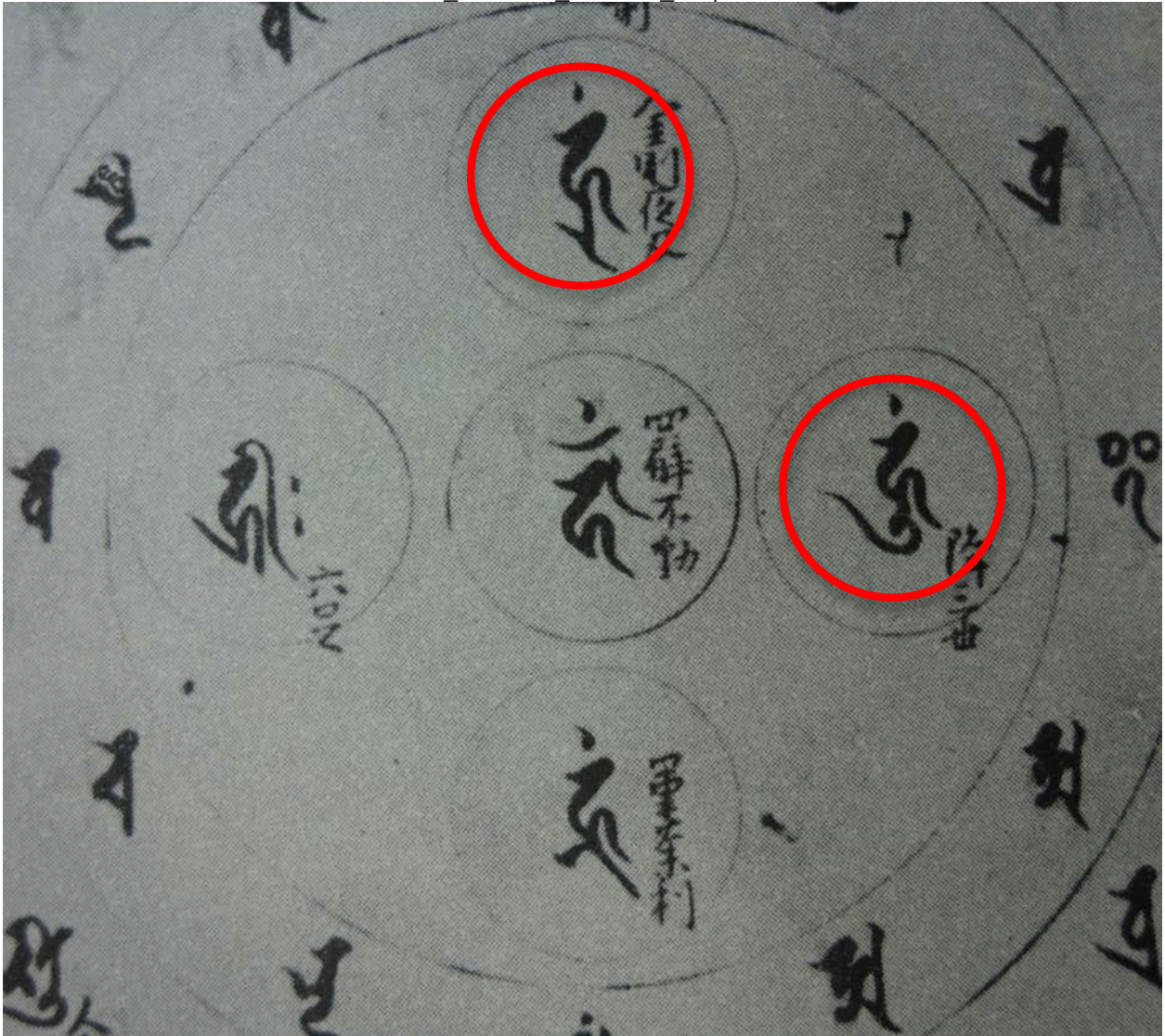
Figure 13. SIDDHAM VOWEL SIGN U are semantically different.
(from 『大正新脩大藏經 図像部』 (Taisho-Shinshu-Daizokyo Zuzoubu), vol. 5, p. 741.)
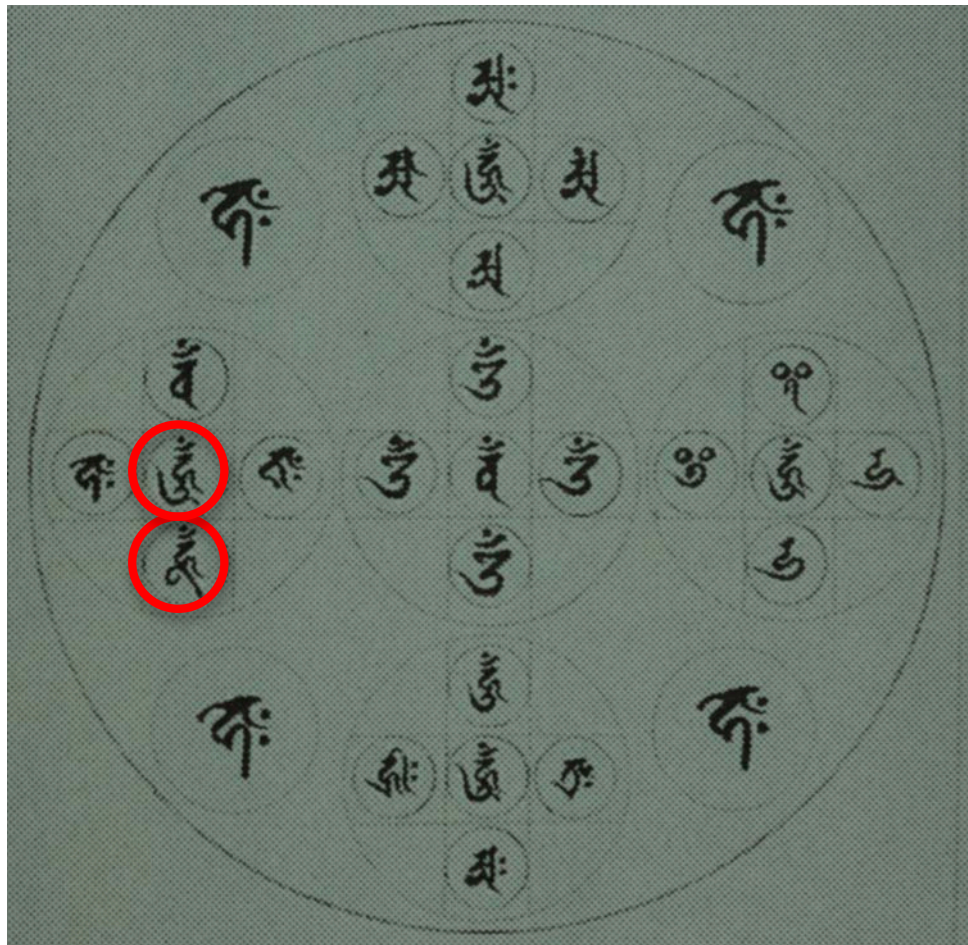
Figure 14. SIDDHAM VOWEL SIGN UU are semantically different.
(from [Kodama: 2002], p. 105.)
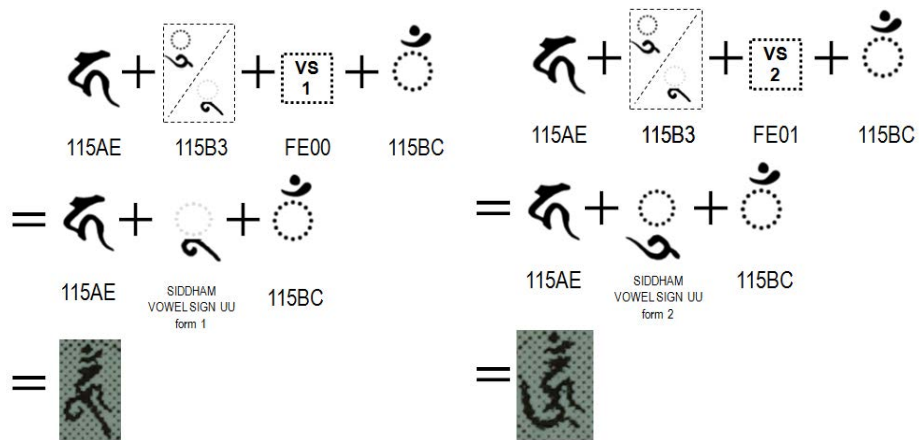
(end of document)

<div style="border:1px solid">

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
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*.**

</div>

A. Administrative

1. **Title:**   *Proposal to encode Variants for Siddham Script*
2. Requester's name: *SAT Project, Kiyonori Nagasaki*
3. Requester type (Member body/Liaison/Individual contribution):   *Individual contribution*
4. Submission date: *2013-04-20*
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):   *No*
      Proposed name of script:
   b. The proposal is for addition of character(s) to an existing block: *Partially Yes*
      Name of the existing block: *Siddham Script*
2. Number of characters in proposal: *7*
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary   **X**   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?
   a. If YES, are the names in accordance with the "character naming guidelines"
      in Annex L of P&P document?   *Yes (variants)*
   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?
      SAT Project
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
      Taichi Kawabata (NTT)
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)?   *No*

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: N4102-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, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? _Yes_
   If YES explain _Variants needed for specific region (Japan)_
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? _SAT Committee, Indic script experts_
   If YES, available relevant documents:
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_
   If YES, where? 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? _Yes_
   If YES, is a rationale for its inclusion provided?
   If YES, reference: _Proposal for Use with Variatin Selector_
9. Can any of the proposed characters be encoded using a composed character sequence of either
   existing characters or other proposed characters? _Yes_
   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? _No_
    If YES, is a rationale for its inclusion provided?
    If YES, reference:
11. Does the proposal include use of combining characters and/or use of composite sequences? _Yes_
    If YES, is a rationale for such use provided? _Yes_
    If YES, reference: _Dependent vowel signs._
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
    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)
    Variation Selectors

13. Does the proposal contain any Ideographic compatibility characters? _No_
    If YES, are the equivalent corresponding unified ideographic characters identified?
    If YES, reference: