Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: **Working Group Document**
Title: **Proposal for encoding seven additional Myanmar characters for Shan Pali in the UCS**
Source: **Michael Everson**
Status: **Individual Contribution**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Date: **2010-09-21**

A number of Myanmar letters used in Buddhist contexts in Shan State in the Union of Myanmar have recently come to light. These were provided by the Rev. Sukhaminda, General Benefactor, and Head of the Shan State Sangha Committee concerned with translating the Tipitaka into Shan, who provided the examples in the figures below, all from books published a few years ago.

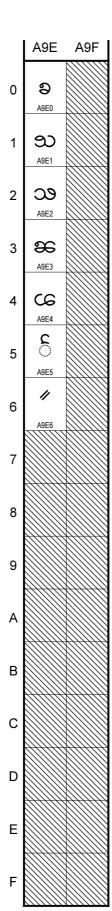The full set of characters used in Shan State to write Pali are as follows:

ဢ *ka*	ခ *kha*	ဢ *ga*	ꨀ *gha*	င *ṅa*	စ *ca*	ꨁ *cha*	ၿ *ja*	ꨂ *jha*	ည *ña*
1075	1076	1077	A9E0	1004	1078	A9E1	AA61	A9E2	107A

ꩦ *ṭa*	ꩧ *ṭha*	ꩨ *ḍa*	ꩩ *ḍha*	ꨃ *ṇa*	တ *ta*	ထ *tha*	ဒ *da*	ဓ *dha*	န *na*
AA66	AA67	AA68	AA69	A9E3	1010	1011	107B	AA6A	107C

ပ *pa*	ဖ *pha*	ဝ *ba*	ꨄ *bha*	မ *ma*	ယ *ya*	꩹ *ra*	လ *la*	ဝ *wa*	သ *sa*
1015	107D	107F	A9E4	1019	101A	AA79	101C	101D	101E

ဟ *ha*	ꩮ *ḷa*	အ *a*
1081	AA6E	1022

Two additional characters are required. The second of these is a reduplication marker indicating that the syllable preceding it is to be pronounced twice. The first is used in Shan only together with the letter WA to indicate a measurement for grain. In the Tai Laing language, however, this same sign is used with a number of base letters.

Unicode Character Properties

```
A9E0;MYANMAR LETTER SHAN GHA;Lo;0;L;;;;;N;;;;;
A9E1;MYANMAR LETTER SHAN CHA;Lo;0;L;;;;;N;;;;;
A9E2;MYANMAR LETTER SHAN JHA;Lo;0;L;;;;;N;;;;;
A9E3;MYANMAR LETTER SHAN NNA;Lo;0;L;;;;;N;;;;;
A9E4;MYANMAR LETTER SHAN BHA;Lo;0;L;;;;;N;;;;;
A9E5;MYANMAR SIGN SHAN SAW;Mc;0;L;;;;;N;;;;;
A9E6;MYANMAR MODIFIER LETTER SHAN REDUPLICATION;Lm;0;L;;;;;N;;;;;
```

	A9E	A9F
0	ၠ A9E0	
1	ၡ A9E1	
2	ၢ A9E2	
3	ၣ A9E3	
4	ၤ A9E4	
5	ၥ A9E5	
6	⁄⁄ A9E6	
7		
8		
9		
A		
B		
C		
D		
E		
F		

Shan Pali consonants

A9E0   ၠ   MYANMAR LETTER KHAMTI GHA
A9E1   ၡ   MYANMAR LETTER KHAMTI CHA
A9E2   ၢ   MYANMAR LETTER KHAMTI JHA
A9E3   ၣ   MYANMAR LETTER KHAMTI NNA
A9E4   ၤ   MYANMAR LETTER KHAMTI BHA

Combining mark

A9E5   ၥ   MYANMAR SIGN SHAN SAW

Punctuation

A9E6   ⁄   MYANMAR MODIFIER LETTER SHAN
               REDUPLICATION

Figure 1. Sample showing ခ GHA.
From page-10 of *The Senghkamleng* Magazine No.11, 2003.

Figure 2. Sample showing ဆ CHA.
From page-100 of *The Senghkamleng* Magazine No.11, 2003.
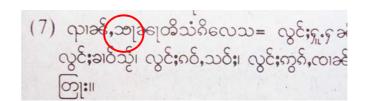
Figure 3. Sample showing ဈ JHA.
From page-65 of *The Senghkamleng* Magazine No.11, 2003.

Figure 4. Sample showing �náNNA.
From page-11 of *The Senghkamleng* Magazine No.11, 2003.

Figure 5. Sample showing ဖ BHA.
From page-41 of *The Senghkamleng* Magazine No.11, 2003.

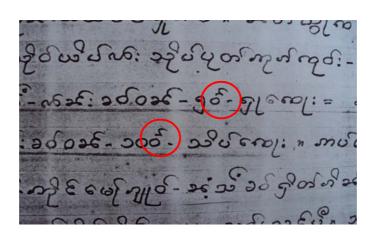Figure 6. Sample showing the ဝ် SAW
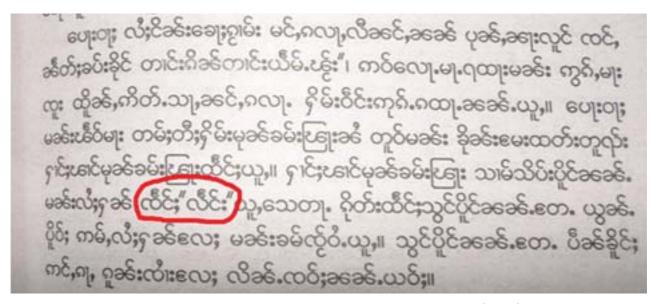From *the Chronicle of Umpawng Hsipaw* by U Mu, the governor of Loilong. (Handwritten)

Figure 7. Sample text showing SHAN REDUPLICATION MARK. The text reads ꟷꟷ, *jeng jeng leng leng*. From page 57 of the Dra Buddhaphasar (Level 5-6-7), 2004.

A. Administrative

1. Title

Proposal for encoding seven additional Myanmar characters for Shan Pali in the UCS.

2. Requester's name

Michael Everson

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

Individual contribution.

4. Submission date

2010-09-21

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

Yes.

6b. More information will be provided later

No.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

No.

Proposed name of script

1b. The proposal is for addition of character(s) to an existing block

Yes.

1c. Name of the existing block

Myanmar Extended-A.

2. Number of characters in proposal

1

3. Proposed category (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)

Category A.

4. Is a repertoire including character names provided?

Yes.

4a. If YES, are the names in accordance with the "character naming guidelines"

Yes.

4b. Are the character shapes attached in a legible form suitable for review?

Yes.

5. Fonts related:

5a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?

Michael Everson.

5b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):

Michael Everson, Fontographer.

6. References:

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

No.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

7. Special encoding issue: 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. Unicode character properties above.

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.

See above.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain

No.

2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

Yes. The Rev. Sukhaminda, General Benefactor, The Head of Shan State Sangha Committee of Translating Tipitaka into Shan.

2a. 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?

People in Myanmar.

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

Common.

4b. Reference

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

Yes.

5b. If YES, where?

In Myanmar.

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

Yes.

6a. If YES, is a rationale provided?

Yes.

6b. If YES, reference

Contemporary use and accordance with the Roadmap.

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

Yes.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

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

Yes.

11b. If YES, is a rationale for such use provided?

No.

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

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

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?