

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 Mon and S’gaw Karen characters in the UCS****Source: Michael Everson and Martin Hosken****Status: Individual Contribution****Replaces: N2768, N1883R; L2/05-178****Action: For consideration by JTC1/SC2/WG2 and UTC****Date: 2006-04-08**

Since the Myanmar script was first encoded, it has been known that a number of additions used by minority languages would be needed. This proposal requests the addition of characters for two of those languages, Mon and S’gaw Karen. It contains the proposal summary form.

Most of the characters proposed here are spacing letters, but combining vowel signs, combining medial consonant signs, and combining tone marks are also proposed.

The history of the Myanmar script is not one of a single line of development. The Mon script, in fact, is probably the oldest of the traditions, though the letterforms of the former have been “Burmicized” due to later influence, particularly that of typography. A number of language-specific differences arose during the period of development, much as has happened with the Arabic and Cyrillic scripts. Most of the letters are used in common, but some letters have language-specific forms. These are not unifiable with “standard” Myanmar letters. For instance, books in Burmese about Mon naturally use both of them concurrently. In the discussion of the additions below, the language-specific letters are listed, in the brief shorthand “x contrasts with Burmese y”.

The new rendering model proposed in N3043 which disunifies visible ASAT from stacking VIRAMA applies to the characters here. The general implication of this is that when following VIRAMA , the LETTERS here should be rendered below and usually slightly smaller than the base letter. Where there is a departure from the mainstream stacking behaviour for Burmese, characters have been added here to facilitate minority-language usage. For example, LETTER NA and LETTER MA subjoin as expected in Burmese NA MA so the special Mon subjoined forms $\text{CONSONANT SIGN MON MEDIAL NA}$ and $\text{CONSONANT SIGN MON MEDIAL MA}$ are encoded as dependent consonant signs.

Additions for Mon

A number of characters contrast with Burmese characters: LETTER MON E contrasts with Burmese LETTER E ; LETTER MON NGA contrasts with Burmese LETTER NGA ; LETTER MON JHA contrasts with Burmese LETTER JHA ; VOWEL SIGN MON II contrasts with Burmese VOWEL SIGN II . Other characters are unique to Mon: VOWEL SIGN MON O represents *open o*; LETTER MON BBA and LETTER MON BBE represent bilabial implosives; $\text{CONSONANT SIGN MON MEDIAL NA}$ and $\text{CONSONANT SIGN MON MEDIAL MA}$ are mentioned above; $\text{CONSONANT SIGN MON MEDIAL LA}$ is used in S’gaw Karen for *ya*, where it contrasts with Burmese $\text{CONSONANT SIGN MEDIAL YA}$ proposed in N3043. A clarification should be given regarding LETTER MON NGA which merits discussion. While Burmese LETTER NGA stacks as expected—

ꠏ—Mon ꠑ LETTER NGA stacks by rendering only its “diacritic” ꠏ; that is, *ꠏ does not occur. It could be possible to treat the diacritic as a medial (like ꠑ CONSONANT SIGN MON MEDIAL NA and ꠒ CONSONANT SIGN MON MEDIAL MA) but this is impractical due to attested glyph variants of the letter. Both ꠑ and ꠒ are attested glyph variants of the letter, and if *ꠏ were encoded as a medial, it could be confusing to users as to whether it should be applied to ꠑ LETTER NGA or to ꠓ VOWEL SIGN E (which isn’t an independent letter, but a dependent vowel). The encoding proposed here is simpler. (Figures 1, 2, 3, 3a, 6, and 8.)

Additions for S’gaw Karen

Four characters used in S’gaw Karen are not used in Burmese: ꠔ LETTER SGAW KAREN SHA contrasts with Burmese ꠕ *rha* (which is ꠖ LETTER RA + the ꠗ CONSONANT SIGN MEDIAL HA proposed in N3043); ꠘ VOWEL SIGN SGAW KAREN EU serves as a vowel [ɣ] and (with the ꠙ SIGN ASAT proposed in N3043) as a tone mark ꠙ; ꠚ SIGN SGAW KAREN HATHI and ꠛ SIGN SGAW KAREN KE PHO are tone marks. The former only occurs with the ꠙ SIGN ASAT, thus ꠛꠙ. It should also be noted that in S’gaw Karen the ꠜ VOWEL SIGN TALL AA proposed in N3043 is used as a vowel (in all contexts following all consonants), and ꠝ VOWEL SIGN AA is used as a tone mark, only occurring with the ꠙ SIGN ASAT, thus ꠙꠝ. (Figures 4, 5, 6, 7, 8, 9, and 10.)

Ordering

The unified order for the Myanmar script incorporating the characters here is given below. Ordering is syllable-based, so this is indicative of only one level of ordering.

ka < kha < ga < gha < nga < mon-nga <
ca < cha < ja < jha < mon-jha < sgaw-karen-sha < nya < nnya <
tta < ttha < dda < ddha < nna <
ta < tha < da < dha < na <
pa < pha < ba < bha < ma <
ya < ra < la < wa < sha < ssa < sa < great-sa < ha < lla < mon-bba <
a < < i < ii < u < uu < vocalic-r < vocalic-rr < vocalic-l < vocalic-ll < e < mon-bbe < mon-e < o < au

Issues

As noted above, it may be observed that ꠔ LETTER SGAW KAREN SHA looks as though it could be a sequence of ꠖ LETTER RA + the ꠗ CONSONANT SIGN MEDIAL HA proposed in N3043. At a “Workshop on Myanmar Language Processing”, held in Yangon 13-15 February 2006 (cf N3043R), this was discussed at length with the S’gaw Karen experts present. The letter has its own place in the alphabet of the languages which use it (it is also used in Western Pwo Karen). In S’gaw Karen, ꠔ LETTER SGAW KAREN SHA is sorted between ꠕ LETTER JA and ꠘ LETTER NNYA; where the ꠗ CONSONANT SIGN MEDIAL HA is used elsewhere in S’gaw Karen the letter is sorted as expected under the base letter (so ꠔ sorts within ꠕ LETTER CHA). In Western Pwo Karen, it is sorted between ꠘ LETTER NNYA and ꠙ LETTER TA. In both languages, the base letter ꠖ *ra* does take a genuine medial (ꠗ *rwa*) which is sorted as expected under ꠖ *ra*. In better typography and in handwriting, ꠔ LETTER SGAW KAREN SHA contrasts with ꠕ *rha* (though much printed matter is not exemplary of the best typography). Since the Myanmar script is to be a unified set to deal with all of these languages, it seems best to let ꠗ CONSONANT SIGN MEDIAL HA be used in its traditional productive medial role in the Burmese, Mon, and S’gaw Karen languages, but to encode as unique letters like ꠔ LETTER SGAW KAREN SHA. Precedent for this can be found in the letterforms used for *jha*: In Burmese, alongside ꠕ is found the form ꠕ, which looks as though it is a ligature of ꠕ *ca* and ꠕ *-ya* (but it isn’t), and in Mon the forms ꠕ and ꠕ are often found, probably better rendered ꠕ in a generic font; the letter is not ꠕ *ja* (+ ꠗ *-ha*) + ꠕ *-ya*.

Unicode Character Properties

```
1028;MYANMAR LETTER MON E;Lo;0;L;;;;N;;;;;
1033;MYANMAR VOWEL SIGN MON II;Mn;0;NSM;;;;;N;;;;;
1034;MYANMAR VOWEL SIGN MON O;Mn;0;NSM;;;;;N;;;;;
105A;MYANMAR LETTER MON NGA;Lo;0;L;;;;N;;;;;
105B;MYANMAR LETTER MON JHA;Lo;0;L;;;;N;;;;;
```

105C;MYANMAR LETTER MON BBA;Lo;0;L;;;;N;;;;;
 105D;MYANMAR LETTER MON BBE;Lo;0;L;;;;N;;;;;
 105E;MYANMAR CONSONANT SIGN MON MEDIAL NA;Mn;0;NSM;;;;N;;;;;
 105F;MYANMAR CONSONANT SIGN MON MEDIAL MA;Mn;0;NSM;;;;N;;;;;
 1060;MYANMAR CONSONANT SIGN MON MEDIAL LA;Mn;0;NSM;;;;N;;;;;
 1061;MYANMAR LETTER SGAW KAREN SHA;Lo;0;L;;;;N;;;;;
 1062;MYANMAR VOWEL SIGN SGAW KAREN EU;Mc;0;L;;;;N;;;;;
 1063;MYANMAR SIGN SGAW KAREN HATHI;Mc;0;L;;;;N;;;;;
 1064;MYANMAR SIGN SGAW KAREN KE PHO;Mc;0;L;;;;N;;;;;

Bibliography

Halliday, R. 1955. *A Mon-English dictionary*. Rangoon: Mon Cultural Section, Ministry of Union Culture, Government of the Union of Burma. (Repring of the 1922 edtion published by the Siam Society, Bangkok.)

Nai Pan Hla. 1989. *An introduction to Mon language*. Kyoto: Center for Southeast Asian Studies, Kyoto University.

Nai Tun Way. 1997. The modern English-Mon dictionary: ပြက္ခိတ်အဘိဓာန် အင်္ဂလိတ်-မန်. Bangkok. The Open Society Insititute Burma Project.

မွန်ဘာသာ: သင်ကြားနည်း (Mon grammar book).

ပုထိုးရှိုဆာချဲ့ ဒဲ ပယုအဒါပယီ = ပိုးကရင်-မြန်မာအဘိဓါန် = *Pwo Kayin-Myanmar Dictionary*. 1989. (Western Pwo Karen).

Acknowledgements

This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley), and also by support from Payap University, Chiang Mai.

Figures

12 MON VOWELS

သွဂ်မန် (၁၂) အခပ်၊

အ	အာ	အု	အူ	ဥ	ဥါ
ဇ	အဲ	အြ	အံ	အံ	အး

VOWEL-SYMBOLS

လကျန်သွဂ်

-	-၁	-၀	-၉	-ါ	-ါါ
-	-`	-၁	-´	-°	-း

Figure 1. Sample from Nai Pan Hla's 1989 Mon dictionary, showing LETTER MON E, VOWEL SIGN MON II, and VOWEL SIGN MON O.

35 MON CONSONANTS

ခန်ကျော် (၂) လွင်ချန်

က	ခ	ဂ	ဃ	င
စ	ဆ	ဇ	ည	ဉ
ဋ	ဌ	ဍ	ဎ	ဏ
တ	ထ	ဒ	ဓ	န
ပ	ဖ	ဗ	ဘ	မ
ယ	ရ	လ	ဝ	သ
ဟ	ဠ	အ	အ	ဢ

Figure 2. Sample from Nai Pan Hla's 1989 Mon dictionary, showing LETTER MON NGA, LETTER MON JHA, LETTER MON BBA, and LETTER MON BBE.

Subscript Letters	Symbols	with combination
င	၇	တူ tūa/təŋa°/
ည	ည	ကူ kŋa/kəŋa°/
ဒ	ဒ	တူ tda/təda°/
န	န	တူ tna/təna°/
မ	မ	တူ tma/təma°/
ယ	ယ	ကျ kya/ca°/
ရ	ရ	ကြ kra/kra°/
လ	လ	ကူ kla/kla°/
ဝ	ဝ	ကူ kwa/kwa°/
စ	စ	ကူ kḥa/kḥa°/
ဟ	ဟ	လှ lha/hla°/

Figure 3. Sample from Nai Pan Hla's 1989 Mon dictionary, showing CONSONANT SIGN MON MEDIAL NA, CONSONANT SIGN MON MEDIAL MA, and CONSONANT SIGN MON MEDIAL LA. In the first line the incorrect form for LETTER MON NGA is shown; elsewhere in the book it is regularly shown correctly as in Figure 2 and in Figure 3a here:

က ခ ဂ ဃ င
ka/ka°/; kha/kha°/; ga/kè°/; gha/khè°/; ṅa/ŋe°/;

လံာ်ပျာ်မိၣ်ပွၣ်တဖၣ်မိအရဲၣ်လီၤတ့ၢ်လံာ်အသး

က	ခ	ဂ	ဃ	င
ကးဒိၣ်ဆဲး	ခးကဝိၤ	ဂးဆဲး	ဃးဒိၣ်	ငးဒီၣ်ရိၣ်
စ	ဆ	ဇ	ဈ	ည
စးပျာ်	ဆးတံ	ဇးသ့ၣ်မး	ဈးခွဲၣ်ပံာ်	ညးညိၣ်ဒိၣ်
တ	ထ	ဒ	န	ပ
တးတံာ်ရ	ထးတဆံာ်	ဒးတဒါ	နးဆဲး	ပးတဝိ
ဖ	ဘ	မ	ယ	ရ
ဖးပဝံာ်	ဘးတကူး	မးခွံးသး	ယးဒီၣ်ထီ	ရးက့ၣ်ကျိၤ
လ	ဝ	သ	ဟ	
လးက့ၣ်ထီၣ်	ဝးကဝိၤ	သးခံာ်	ဟးက့ၣ်လီၤ	
	အ	ဇ		
	အးခံာ်	ဇးကတၢ်		

Figure 4. Sample from a reader in S'gaw Karen, showing LETTER SGAW KAREN SHA.

လိာ်ဖျါနိာ်ပိာ်တၢ်က့ၤက့ၤလၢအညီဆူအကီ
 ပီး
 တၢ်က့ၤအက့ၤ

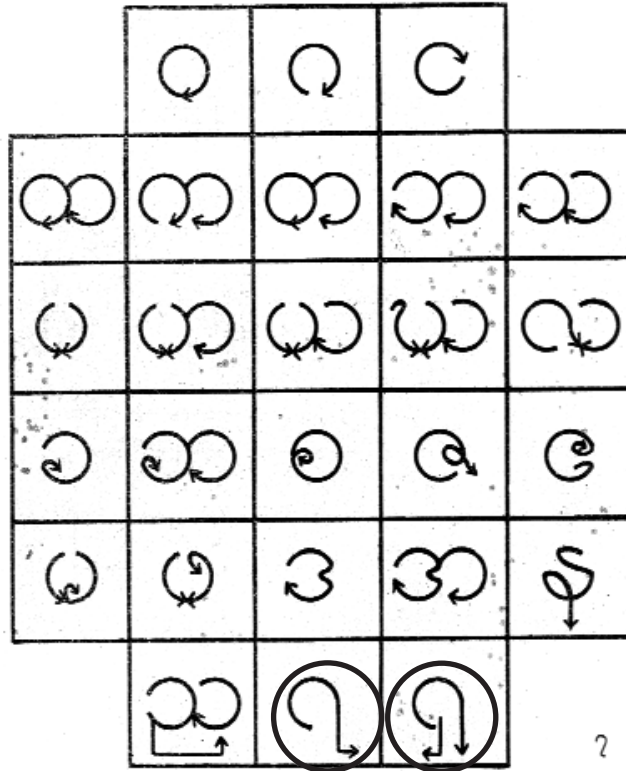


Figure 5. Sample from a reader in S'gaw Karen, showing the ductus for writing S'gaw Karen letters. Note the difference between LETTER RA and LETTER SGAW KAREN SHA.

လိာ်ဖျါနိာ်တၢ်ပာ်

၂ ပုၤဂ် အစၢ်စးတၢ်ခိးလိာ်ဖျါနိာ်အဂၢၢ်အံၤ - ဝ + ဂ = ငှ | ဝ + ဂ = ဟ

၃ " ယ " " " " " - ဝ + ယ = ဝှ | က + ယ = ကှ

၄ " ခ " " " " " = က + ခ = ကျ | ဝ + ခ = ဝှ

၅ " ဓ " " " " " - က + ဓ = ကျ | ဓ + ဓ = ဝှ

၆ " ဝ " " " " " = က + ဝ = ကှ | ဝ + ဝ = ဝှ

၇ မုၢ်လိာ်ဖျါနိာ်တၢ်ပာ်စးပီးဝဲက့ၤတီၢ်ခိးလိာ်ဖျါနိာ် အဖျါနိာ်တၢ်ခိးလိာ်.
 ဝဲစိ ၇ + ၇ = ၁၄ (နိာ်စိပတြာ်--(Geometry)

Figure 6. Sample from a reader in S'gaw Karen, showing CONSONANT SIGN MON MEDIAL LA (used in S'gaw Karen for ya) alongside characters proposed in N3043, CONSONANT SIGN MEDIAL HA (used in S'gaw Karen for ga), CONSONANT SIGN MEDIAL RA, CONSONANT SIGN MEDIAL YA (used in S'gaw Karen for la), and CONSONANT SIGN MEDIAL WA.

တၢ်သီၣ်ခိၣ်ပုၣ်

၁.	-ါ	=	အါ	←	အး	အါ	=	အါ
၂.	-ံ	=	အံ		အး	အံ	=	အံ
၃.	-ါ	=	အါ	←	အး	အါ	=	အါ
၄.	-ါ	=	အါ		အး	အါ	=	အါ
၅.	-ါ	=	အါ		အး	အါ	=	အါ
၆.	-ါ	=	အါ		အး	အါ	=	အါ
၇.	-ါ	=	အါ		အး	အါ	=	အါ
၈.	-ါ	=	အါ		အး	အါ	=	အါ
၉.	-ါ	=	အါ		အး	အါ	=	အါ

တၢ်သီၣ်ထိးနီၤ

၁.	-ါ	=	အါသံ	အး	အါ	အါ	သံ	=	အါ
၂.	-ံ	=	အါသံ	အး	အါ	အါ	သံ	=	အါ
၃.	-ါ	=	အါသံ	အး	အါ	အါ	သံ	=	အါ
၄.	-ါ	=	အါသံ	အး	အါ	အါ	သံ	=	အါ
၅.	-ါ	=	အါသံ	အး	အါ	အါ	သံ	=	အါ
၆.	-ါ	=	အါသံ	အး	အါ	အါ	သံ	=	အါ
၇.	-ါ	=	အါသံ	အး	အါ	အါ	သံ	=	အါ
၈.	-ါ	=	အါသံ	အး	အါ	အါ	သံ	=	အါ
၉.	-ါ	=	အါသံ	အး	အါ	အါ	သံ	=	အါ
၁၀.	-ါ	=	အါသံ	အး	အါ	အါ	သံ	=	အါ
၁၁.	-ါ	=	အါသံ	အး	အါ	အါ	သံ	=	အါ

Figure 7. Sample from a reader in S'gaw Karen, showing vowels and tone marks, including VOWEL SIGN TALL AA (proposed in N3043), VOWEL SIGN SGAW KAREN EU, SIGN SGAW KAREN HATHI, SIGN SGAW KAREN KE PHO, and CONSONANT SIGN MON MEDIAL LA.

ပနၢ်ဝါ

ခိၣ် ဝိၣ် မိၣ် လီၤ ပနၢ် ဝါ	လဲၤ ပဲၣ် ဘၣ် လၢ ဆိၣ် ခိၣ် သ့ၣ်
ဟ် ဟ် မိၣ် လီၤ ပနၢ် ဝါ	ဆိၣ် ခိၣ် သ့ၣ် ပှၤ ဝံ လၢ ဟိး
မိၣ် လီၤ လၢ ဘ့ၣ် ခိၣ် ဘ့ၣ် လၢ	ထိး ဝါ ကိၣ် ပှၤ ခိၣ် အ ကိၣ်
ဘ့ၣ် ခိၣ် ဘ့ၣ် လၢ လဲၤ မၤ ပှၤ	ခိၣ် အ ကိၣ် လၢ နး တိၣ် စ့
လဲၤ မၤ ဘၣ် စိၣ် ဝံ မိၣ် ခါ	နး တိၣ် စ့ အ က ဝါ တိၣ်
စိၣ် ဝံ မိၣ် ခါ စိၣ် ခါ နၢ်	အ မဲၣ် ထံ လီၤ စိၣ် လီၤ စိၣ်
စိၣ် ခါ နၢ် ဝံ ကျိၣ် ဝံ ကျိၣ်	လီၤ စိၣ် ဘၣ် ဆိၣ် ခိၣ် တ ကိၣ်
ဆး ထူ မိၣ် ဘၣ် ပိၣ် ပိၣ် ကျိၣ်	ဆိၣ် ခိၣ် တ ကိၣ် လံၣ် လူၣ် ဘိ
ပိၣ် ပိၣ် ကျိၣ် လဲၤ ပဲၣ် အ ဝါ	လံၣ် လူၣ် ဘိ ကဲ ထီၣ် လၢ နီ
	ကဲ ထီၣ် နီ ပှၤ က ညိၣ် ဘိၣ်

Figure 8. Sample from a reader in S'gaw Karen, showing SIGN SGAW KAREN HATHI and SIGN SGAW KAREN KE PHO. The CONSONANT SIGN MEDIAL WA (proposed in N3043) is shown with a triangular glyph.

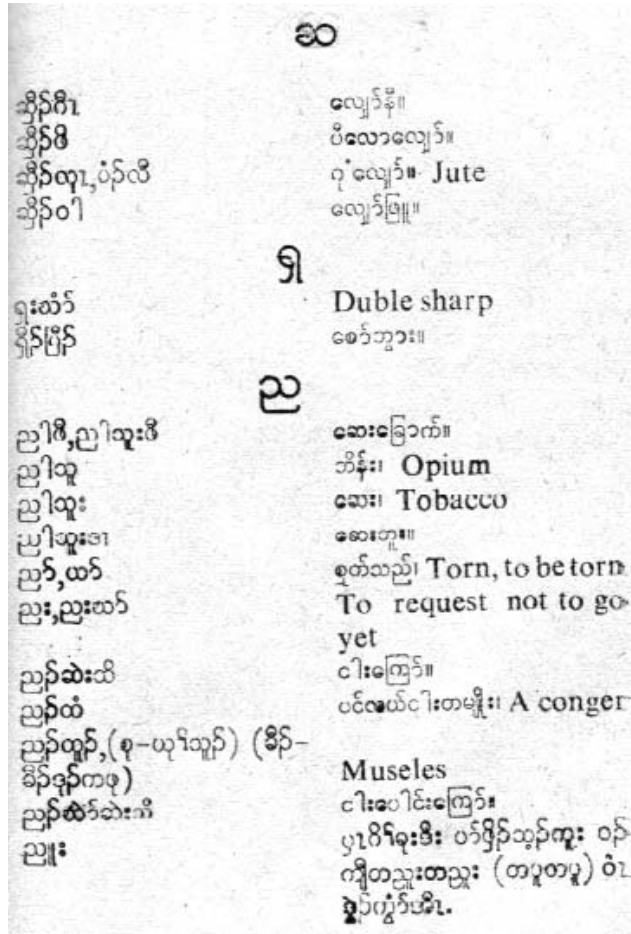


Figure 9. Sample from a short dictionary of S'gaw Karen, showing LETTER SGAW KAREN SHA in its position between LETTER CHA and LETTER NNYA.

က	ခ	ဂ	ဃ	င
စ	ဆ	ဇ	ည	ရှ
ဒ	ဇ	ဒ	န	ပ
ဖ	တ	မ	ယ	ရ
လ	ဝ	ဝ	တ	အ
		ဇ		

က	ခ	ဂ	ဃ	င
စ	ဆ	ဇ	ည	ရှ
တ	ထ	ဒ	န	ပ
ဖ	တ	မ	ယ	ရ
လ	ဝ	ဝ	တ	အ
		ဇ	ပ	

Figure 10. Two samples from Western Pwo Karen primers, showing LETTER SGAW KAREN SHA in its position between LETTER NNYA and LETTER TA.

TABLE XX - Row 10: MYANMAR

	100	101	102	103	104	105	106	107	108	109
0	၀	၁	၂	၃	၄	၅	၆			
1	၇	၈	၉	၁၀	၁၁	၁၂	၁၃			
2	၁၄	၁၅		၁၆	၁၇	၁၈	၁၉			
3	၂၀	၂၁	၂၂	၂၃	၂၄	၂၅	၂၆			
4	၂၇	၂၈	၂၉	၃၀	၃၁	၃၂	၃၃			
5	၃၄	၃၅	၃၆		၃၇	၃၈				
6	၃၉	၄၀	၄၁	၄၂	၄၃	၄၄				
7	၄၅	၄၆	၄၇	၄၈	၄၉	၅၀				
8	၅၁	၅၂	၅၃	၅၄	၅၅	၅၆				
9	၅၇	၅၈	၅၉	၆၀	၆၁	၆၂				
A	၆၃	၆၄	၆၅	၆၆	၆၇	၆၈				
B	၆၉	၇၀	၇၁	၇၂	၇၃	၇၄				
C	၇၅	၇၆	၇၇	၇၈	၇၉	၈၀				
D	၈၁	၈၂	၈၃	၈၄	၈၅	၈၆				
E	၈၇	၈၈	၈၉	၉၀	၉၁	၉၂				
F	၉၃	၉၄	၉၅	၉၆	၉၇	၉၈				

G = 00
P = 00

TABLE XX - Row 10: MYANMAR

hex	Name
00	MYANMAR LETTER KA
01	MYANMAR LETTER KHA
02	MYANMAR LETTER GA
03	MYANMAR LETTER GHA
04	MYANMAR LETTER NGA
05	MYANMAR LETTER CA
06	MYANMAR LETTER CHA
07	MYANMAR LETTER JA
08	MYANMAR LETTER JHA
09	MYANMAR LETTER NYA
0A	MYANMAR LETTER NNYA
0B	MYANMAR LETTER TTA
0C	MYANMAR LETTER TTHA
0D	MYANMAR LETTER DDA
0E	MYANMAR LETTER DDHA
0F	MYANMAR LETTER NNA
10	MYANMAR LETTER TA
11	MYANMAR LETTER THA
12	MYANMAR LETTER DA
13	MYANMAR LETTER DHA
14	MYANMAR LETTER NA
15	MYANMAR LETTER PA
16	MYANMAR LETTER PHA
17	MYANMAR LETTER BA
18	MYANMAR LETTER BHA
19	MYANMAR LETTER MA
1A	MYANMAR LETTER YA
1B	MYANMAR LETTER RA
1C	MYANMAR LETTER LA
1D	MYANMAR LETTER WA
1E	MYANMAR LETTER SA
1F	MYANMAR LETTER HA
20	MYANMAR LETTER LLA
21	MYANMAR LETTER A
22	(This position shall not be used)
23	MYANMAR LETTER I
24	MYANMAR LETTER II
25	MYANMAR LETTER U
26	MYANMAR LETTER UU
27	MYANMAR LETTER E
28	MYANMAR LETTER <u>MON E</u>
29	MYANMAR LETTER O
2A	MYANMAR LETTER AU
2B	MYANMAR VOWEL SIGN TALL AA
2C	MYANMAR VOWEL SIGN AA
2D	MYANMAR VOWEL SIGN I
2E	MYANMAR VOWEL SIGN II
2F	MYANMAR VOWEL SIGN U
30	MYANMAR VOWEL SIGN UU
31	MYANMAR VOWEL SIGN E
32	MYANMAR VOWEL SIGN AI
33	MYANMAR VOWEL SIGN <u>MON II</u>
34	MYANMAR VOWEL SIGN <u>MON O</u>
35	(This position shall not be used)
36	MYANMAR SIGN ANUSVARA
37	MYANMAR SIGN DOT BELOW
38	MYANMAR SIGN VISARGA
39	MYANMAR SIGN VIRAMA
3A	MYANMAR SIGN ASAT
3B	MYANMAR CONSONANT SIGN MEDIAL YA
3C	MYANMAR CONSONANT SIGN MEDIAL RA
3D	MYANMAR CONSONANT SIGN MEDIAL WA
3E	MYANMAR CONSONANT SIGN MEDIAL HA
3F	MYANMAR LETTER GREAT SA
40	MYANMAR DIGIT ZERO
41	MYANMAR DIGIT ONE
42	MYANMAR DIGIT TWO
43	MYANMAR DIGIT THREE
44	MYANMAR DIGIT FOUR
45	MYANMAR DIGIT FIVE
46	MYANMAR DIGIT SIX
47	MYANMAR DIGIT SEVEN
48	MYANMAR DIGIT EIGHT
49	MYANMAR DIGIT NINE
4A	MYANMAR SIGN LITTLE SECTION
4B	MYANMAR SIGN SECTION
4C	MYANMAR SYMBOL LOCATIVE
4D	MYANMAR SYMBOL COMPLETED
4E	MYANMAR SYMBOL AFOREMENTIONED
4F	MYANMAR SYMBOL GENITIVE
50	MYANMAR LETTER SHA
51	MYANMAR LETTER SSA
52	MYANMAR LETTER VOCALIC R
53	MYANMAR LETTER VOCALIC RR
54	MYANMAR LETTER VOCALIC L
55	MYANMAR LETTER VOCALIC LL
56	MYANMAR VOWEL SIGN VOCALIC R
57	MYANMAR VOWEL SIGN VOCALIC RR
58	MYANMAR VOWEL SIGN VOCALIC L

hex	Name
59	MYANMAR VOWEL SIGN VOCALIC LL
5A	MYANMAR LETTER MON NGA
5B	MYANMAR LETTER MON JHA
5C	MYANMAR LETTER MON BBA
5D	MYANMAR LETTER MON BBE
5E	MYANMAR CONSONANT SIGN MON MEDIAL NA
5F	MYANMAR CONSONANT SIGN MON MEDIAL MA
60	MYANMAR CONSONANT SIGN MON MEDIAL LA
61	MYANMAR LETTER SGAW KAREN SHA
62	MYANMAR LETTER SGAW KAREN EU
63	MYANMAR SIGN SGAW KAREN HATHI
64	MYANMAR SIGN SGAW KAREN KE PHO
65	(This position shall not be used)
66	(This position shall not be used)
67	(This position shall not be used)
68	(This position shall not be used)
69	(This position shall not be used)
6A	(This position shall not be used)
6B	(This position shall not be used)
6C	(This position shall not be used)
6D	(This position shall not be used)
6E	(This position shall not be used)
6F	(This position shall not be used)
70	(This position shall not be used)
71	(This position shall not be used)
72	(This position shall not be used)
73	(This position shall not be used)
74	(This position shall not be used)
75	(This position shall not be used)
76	(This position shall not be used)
77	(This position shall not be used)
78	(This position shall not be used)
79	(This position shall not be used)
7A	(This position shall not be used)
7B	(This position shall not be used)
7C	(This position shall not be used)
7D	(This position shall not be used)
7E	(This position shall not be used)
7F	(This position shall not be used)
80	(This position shall not be used)
81	(This position shall not be used)
82	(This position shall not be used)
83	(This position shall not be used)
84	(This position shall not be used)
85	(This position shall not be used)
86	(This position shall not be used)
87	(This position shall not be used)
88	(This position shall not be used)
89	(This position shall not be used)
8A	(This position shall not be used)
8B	(This position shall not be used)
8C	(This position shall not be used)
8D	(This position shall not be used)
8E	(This position shall not be used)
8F	(This position shall not be used)
90	(This position shall not be used)
91	(This position shall not be used)
92	(This position shall not be used)
93	(This position shall not be used)
94	(This position shall not be used)
95	(This position shall not be used)
96	(This position shall not be used)
97	(This position shall not be used)
98	(This position shall not be used)
99	(This position shall not be used)
9A	(This position shall not be used)
9B	(This position shall not be used)
9C	(This position shall not be used)
9D	(This position shall not be used)
9E	(This position shall not be used)
9F	(This position shall not be used)

A. Administrative

1. Title

Proposal for encoding Mon and S'gaw Karen characters in the UCS.

2. Requester's name

Michael Everson and Martin Hosken

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

Individual contribution.

4. Submission date

2006-04-08

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.

2. Number of characters in proposal

14

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.

4a. Proposed Level of Implementation (1, 2 or 3)

Level 2

4b. Is a rationale provided for the choice?

Yes.

4c. If YES, reference

Brahmic Level 2 implementation.

5a. Is a repertoire including character names provided?

Yes.

5b. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?

Yes.

5c. Are the character shapes attached in a legible form suitable for review?

Yes.

6a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson.

6b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

7a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

No.

7b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

8. 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.

9. 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.

Yes. See N2768.

2a. 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.

2b. If YES, with whom?

San Lwin (Director General, Myanmar Language Commission), **Tun Tint** (Myanmar Language Commission), **Thein Oo** (President, Myanmar Computer Federation), **Kyaw Thein** (Vice-President, Myanmar Computer Federation), **Myint Myint Than** (Director, Myanmar Computer Federation), **Zaw Htut** (Myanmar Computer Professional Association, Myanmar's NET), **Htoo Myint Naung** (Myanmar Project, Technomation Studios, Universities of Computer Studies Yangon), **Myint Thu** (Myanmar Project, Myanmar Heritage Publications), **Ngwe Tun** (Mon—Myanmar Computer Professional Association, Solveware Solution, Myanmar Info-Tech), **Maung Maung Thant** (Myanmar Computer Professional Association), **Jai Pah Bung Mein** (Shan—SSi Technologies), **Saw Hare Sei** (S'gaw Karen—Ayeyawady Data Centre), **Saw Baldwin Khaing Oo** (S'gaw Karen—Ayeyawady Data Centre), **Nant Silver Tun** (Western Pwo Karen—Pwo Kayin Conference), **William Wai Lin Kyaw** (Myanmar Computer Professional Association, Myanmar Linux Users Group), **Ye Myat Thu** (Alpha Mandalay, Alpha Info-Tech), **Keith Stribley** (Thanlwinsoft).

2c. 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.

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

Yes.

6b. If YES, is a rationale provided?

Yes.

6c. 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)?

N/A.

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?

Yes.

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

Yes.

10c. If YES, reference

See "Issues" above.

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?

Yes.

11c. If YES, reference

Brahmic vowel and consonant signs.

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?