

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 Myanmar characters for Shan and Palaung in the UCS****Source: UC Berkeley Script Encoding Initiative (Universal Scripts Project)****Authors: Michael Everson and Martin Hosken****Status: Liaison Contribution****Replaces: N3080****Action: For consideration by JTC1/SC2/WG2 and UTC****Date: 2006-09-08**

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 a number of them. It contains the proposal summary form.

The languages supported by this proposal are the Tai-Kadai language Shan and the Mon-Khmer language Rumai Palaung. Eleven of the characters proposed are spacing letters, but one combining consonant sign, four combining vowel signs, and seven combining tone marks are also proposed.

The history of the Myanmar script is not one of a single line of development. 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, and books in Burmese about Karen, for instance, 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”.

Additions for Shan

A number of characters contrast with Burmese characters: 𑜀 LETTER SHAN A contrasts with Burmese ာ LETTER A; 𑜁 LETTER SHAN KA contrasts with Burmese ာ LETTER KA; 𑜂 LETTER SHAN KHA contrasts with Burmese ာ LETTER KHA; 𑜃 LETTER SHAN CA contrasts with Burmese ာ LETTER CA; 𑜄 LETTER SHAN NYA contrasts with Burmese ာ LETTER NYA; 𑜅 LETTER SHAN NA contrasts with Burmese ာ LETTER NA; 𑜆 LETTER SHAN PHA contrasts with Burmese ာ LETTER PHA; 𑜇 LETTER SHAN THA contrasts with Burmese ာ LETTER SA; 𑜈 LETTER SHAN HA contrasts with Burmese ာ LETTER HA; 𑜉 VOWEL SIGN SHAN AA contrasts with Burmese ာ VOWEL SIGN AA; 𑜊 CONSONANT SIGN SHAN MEDIAL WA contrasts with the Burmese ာ CONSONANT SIGN MEDIAL WA. Other characters are unique to Shan: 𑜋 LETTER SHAN FA represents [f]; 𑜌 VOWEL SIGN SHAN E represents open e; 𑜍 VOWEL SIGN SHAN E ABOVE represents open e word-internally; 𑜎 VOWEL SIGN SHAN FINAL Y is used in rising diphthongs. Shan extends the VISARGA function with additional tone marks 𑜏, SIGN SHAN TONE-2, 𑜐, SIGN SHAN TONE-3, 𑜑, SIGN SHAN COUNCIL TONE-4 (used in Shan Council orthography), 𑜒, SIGN SHAN TONE-5, 𑜓, SIGN SHAN TONE-6, and 𑜔, SIGN SHAN COUNCIL EMPHATIC TONE (used in Shan Council orthography). (Figures 1, 2, 3, and 4.)

Additions for Rumai Palaung

Rumai Palaung makes use of a unique character: 𑜕 LETTER RUMAI PALAUNG FA. It also uses four tone marks, two of which are already encoded: U+1037 𑜖 MYANMAR SIGN DOT BELOW is used for tone 1, U+1038 𑜗 MYANMAR SIGN VISARGA for tone 3, and three which are proposed here: *U+1083 𑜘, SIGN SHAN TONE-2 (used for tone 4), *U+1084 𑜙, SIGN SHAN TONE-3 (used for tone 5), and 𑜚, SIGN RUMAI PALAUNG TONE-6 are used (Figure 5.)

Ordering

The unified order for the Myanmar script incorporating the characters here (and those previously accepted for encoding) is given below. Ordering is syllable-based, so this is indicative of only one level of ordering. The Karen and Kayah characters proposed in N3xxx are shown here in italics.

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

Issues

Several characters look as though they could be sequences of a base character plus U+103E ၵ CONSONANT SIGN MEDIAL HA. These are like ၵ LETTER SGAW KAREN SHA which has already been accepted for encoding: the letters are ၶ LETTER SHAN FA and ၷ LETTER SHAN HA. At a “Workshop on Myanmar Language Processing”, held in Yangon 13-15 February 2006 (cf N3043R), this was discussed at length. Both ၶ LETTER SHAN FA and ၷ LETTER SHAN HA can be seen with alternate shapes ၶ and ၷ (see Figure 3). It’s probable that the origin of the former is ၶ *pha* + ၵ *-ha*, but the origin of the latter is (according so Sai Kam Mong 2004) a shape like ၶ, where the top part isn’t analysable as any other letter. Since the Myanmar script is to be a unified set to deal with all of these languages, we judge it best to let ၵ be used in its traditional productive medial role in the Burmese, Mon, and S’gaw Karen languages, but to encode as unique letters the ones used non-productively in Shan.

Unicode Character Properties

1022;MYANMAR LETTER SHAN A;Lo;0;L;;;;N;;;;;
1075;MYANMAR LETTER SHAN KA;Lo;0;L;;;;N;;;;;
1076;MYANMAR LETTER SHAN KHA;Lo;0;L;;;;N;;;;;
1077;MYANMAR LETTER SHAN CA;Lo;0;L;;;;N;;;;;
1078;MYANMAR LETTER SHAN NYA;Lo;0;L;;;;N;;;;;
1079;MYANMAR LETTER SHAN NA;Lo;0;L;;;;N;;;;;
107A;MYANMAR LETTER SHAN PHA;Lo;0;L;;;;N;;;;;
107B;MYANMAR LETTER SHAN FA;Lo;0;L;;;;N;;;;;
107C;MYANMAR LETTER SHAN THA;Lo;0;L;;;;N;;;;;
107D;MYANMAR LETTER SHAN HA;Lo;0;L;;;;N;;;;;
107E;MYANMAR CONSONANT SIGN SHAN MEDIAL WA;Mn;0;NSM;;;;N;;;;;
107F;MYANMAR VOWEL SIGN SHAN AA;Mc;0;L;;;;N;;;;;
1080;MYANMAR VOWEL SIGN SHAN E;Mc;0;NSM;;;;N;;;;;
1081;MYANMAR VOWEL SIGN SHAN E ABOVE;Mn;0;NSM;;;;N;;;;;
1082;MYANMAR VOWEL SIGN SHAN FINAL Y;Mn;0;NSM;;;;N;;;;;
1083;MYANMAR SIGN SHAN TONE-2;Mc;0;L;;;;N;;;;;
1084;MYANMAR SIGN SHAN TONE-3;Mc;0;L;;;;N;;;;;
1085;MYANMAR SIGN SHAN COUNCIL TONE-4;Mc;0;L;;;;N;;;;;
1086;MYANMAR SIGN SHAN TONE-5;Mc;0;L;;;;N;;;;;
1087;MYANMAR SIGN SHAN TONE-6;Mc;0;L;;;;N;;;;;
1088;MYANMAR SIGN SHAN COUNCIL EMPHATIC TONE;Mn;0;NSM;;;;N;;;;;
1089;MYANMAR VOWEL SIGN RUMAI PALAUNG FA;Lo;0;L;;;;N;;;;;
108A;MYANMAR SIGN RUMAI PALAUNG TONE-6;Mc;0;L;;;;N;;;;;

Bibliography

Sung Sum, Gant Kham. 2001. ၵပိတ်,ခေ့ၵ်းတၢ်း = *Shan dictionary*. လၢးသီၵ်း [Lashio]: [s.n.].
Sao Tern Moeng. 1995. *Shan-English dictionary*. Kensington: Dunwoody Press. ISBN 0-93-1745-92-6
ခူးပိတ်းလီ [Khū: Mūing: Lī]. [s.d.]. ၵပိတ်းလိၵ်းထံး ၵၢ်ဆံၵ်း (Shan-Thai reader).
သုၵ်းခေ့ၵ်းလိၵ်းထံး 200 ၵပိတ်းပိတ်း (Shan-Thai reader).

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

သွၚ်ခၢ်ကၢၤခၢ်,လိၤဂ်းတံးဂှ်း ဂှ်းလိၤဝ်းမၤမၤလၢ; ပိးခၢ်ဝ်.တံးခၢ်

- ◆ မးသိၣ်ဝ်း
 က ခ င လ သ ဝု တ ထ ခ
 ဝ ဃ ဗု မ ယ ရ လ ဝ ဖု က
- ◆ မးသိၣ်ဂီၢ်
 ဂ တ ဗ သ
- ◆ မးဂပ်းဝ်း
 က ကါ ကိ ကီ ကူ ကူ
 ကေ ငက ကူဝ် ကေၤ ကိုဝ် ကိုဝ်
- ◆ မးသိၣ်သွၚ်.
 ကံ ကံၤ ကူၣ် ကူၣ် ကွံ ကိုၣ် ကိုၣ်
- ◆ မးဂပ်းသွၚ်.
 ကဝ် ကာဝ် ကိဝ် ကီဝ် ကိဝ် ကွံ
- ◆ တုၢ်မိတ်းသိၣ်လၢ
 ကမ် ကာမ် ကိမ် ကီမ် ကိမ် ကုမ် ကူမ် ကွမ် ကိုမ် ကိုမ်
 ကခမ် ကာခမ် ကိခမ် ကီခမ် ကိခမ် ကုခမ် ကူခမ် ကွခမ် ကိုခမ် ကိုခမ်
 ကင် ကာင် ကိင် ကီင် ကိင် ကုင် ကူင် ကွင် ကိုင် ကိုင်
- ◆ တုၢ်မိတ်းသိၣ်ခါတ်.
 ကပ် ကာပ် ကိပ် ကီပ် ကိပ် ကုပ် ကူပ် ကွပ် ကိုပ် ကိုပ်
 ကတ် ကာတ် ကိတ် ကီတ် ကိတ် ကုတ် ကူတ် ကွတ် ကိုတ် ကိုတ်
 ကဂ် ကာဂ် ကိဂ် ကီဂ် ကိဂ် ကုဂ် ကူဂ် ကွဂ် ကိုဂ် ကိုဂ်
- ◆ မးဂပ်းသွၚ်.
 က + ဝ = ကွ က + ယ = ကျ က + ရ = ကြ
- ◆ မံးသိၣ်
 ကါ ကါး ကါး ကါး ကါး ကါး

Figure 1. Sample from a Shan-Thai reader, showing
 က LETTER SHAN KA, ခ LETTER SHAN KHA, င LETTER SHAN CA, ဝု LETTER SHAN NYA, ခ LETTER SHAN NA,
 ဝ LETTER SHAN PHA, ဃ LETTER SHAN FA, ဖု LETTER SHAN HA, တ LETTER SHAN A,
 ထ LETTER SHAN THA, ဝု VOWEL SIGN SHAN E, ဝါ VOWEL SIGN SHAN AA, ဝံ VOWEL SIGN SHAN FINAL Y,
 ဝီ VOWEL SIGN SHAN E ABOVE, ဝု CONSONANT SIGN SHAN MEDIAL WA,
 ဝ, SIGN SHAN TONE-2, ဝး SIGN SHAN TONE-3, ဝ. SIGN SHAN TONE-5, and ဝး SIGN SHAN TONE-6.

လၢရ်းခပုံတူဝ်လိက်;ခဵုးပပုံခရ်းနမ်း

က	ခ	င	ဏ	သ	ယ				
တ	ထ	ဆ	ပ	ဖ	မ				
	လ	ဝ	ှ	ာ					
---	ၵ	ၶ	း	ၵ	း				
က	ကျ	ကီ	ကော	ကော	ကူ	ကူဝ်	ကော့	ကိုဝ်	ကိုဝ်
ကို	-	-	-	-	-	-	-	-	-
ကရ်	ကရ်	-	-	-	ကူရ်	ကူရ်	ကွရ်	ကိုရ်	ကိုရ်
ကဂ်	ကဂ်	ကိဂ်	ကိဂ်	ကိဂ်	ကုဂ်	ကုဂ်	ကွဂ်	ကိုဂ်	ကိုဂ်
ကင်	ကင်	ကိင်	ကိင်	ကိင်	ကုင်	ကုင်	ကွင်	ကိုင်	ကိုင်
ကတ်	ကတ်	ကိတ်	ကိတ်	ကိတ်	ကုတ်	ကုတ်	ကွတ်	ကိုတ်	ကိုတ်
ကခ်	ကခ်	ကိခ်	ကိခ်	ကိခ်	ကုခ်	ကုခ်	ကွခ်	ကိုခ်	ကိုခ်
ကပ်	ကပ်	ကိပ်	ကိပ်	ကိပ်	ကုပ်	ကုပ်	ကွပ်	ကိုပ်	ကိုပ်
ကမ်	ကမ်	ကိမ်	ကိမ်	ကိမ်	ကုပ်	ကုပ်	ကွမ်	ကိုမ်	ကိုမ်
ကပ်	ကပ်	ကိပ်	ကိပ်	ကိပ်	-	-	-	-	-

(၁၁၁)

Figure 2. Sample from a Sung Sum’s Shan dictionary, showing က LETTER SHAN KA, ခ LETTER SHAN KHA, င LETTER SHAN CA, ဆ LETTER SHAN NA, ဖ LETTER SHAN PHA, ဖ LETTER SHAN HA, ဏ LETTER SHAN A, ဵ SIGN SHAN TONE-2, ံ SIGN SHAN TONE-3, ့ SIGN SHAN TONE-5, and း SIGN SHAN TONE-6, ဵ ဝ VOWEL SIGN SHAN AA, ံ ဝ VOWEL SIGN SHAN E, ့ ဝ CONSONANT SIGN SHAN MEDIAL WA, း VOWEL SIGN SHAN E ABOVE,

ဖ — ဖှ

ဖ The eleventh letter of the Shan alphabet, the fourth in the class of aspirated consonants and the second in the class of labials. ဖှ, which is in the class of labiodental, is the twelfth letter of the new Shan alphabet, and the usage of ဖ or ဖှ is determined by the dialect of the region.

Figure 3. Sample from Sao Tern Moeng 1995 showing alternate glyphs for ဖ LETTER SHAN PHA and ဖှ LETTER SHAN FA.

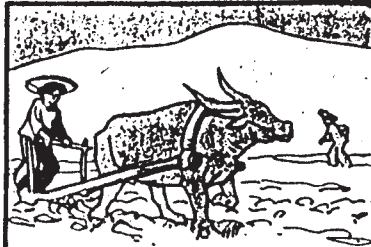

<p> နှံ့သ် နှံ့ တံး နှံး ထံးခါး ဝံးခါး လံးနှံ့ </p>	<p> မါ နှံး ဝံးမါး ဂါဝံး နှံ့ဝံး တီဝံး ဝံးလံ </p>	 <p>ဝံးထံးခါး</p>
 <p>ဝံးလိဝံး</p>	<p> ဝံးမံးနှံ့လံးနှံးခါး ဝံးပေမါးဝံးတီဝံး ဝံးဝံးဝံးတီဝံး ဂါ ဝံးမါးလံးဝံး ပေမါးဝံးယံ့ဝံးယူဝံးလဲဝံး ဝံးဝံးတီဝံးဝံးတံး </p>	

Figure 4. Sample from a Shan Council reader, showing ◌ံ SIGN SHAN COUNCIL EMPHATIC TONE (in the three examples to the right) and ◌ံ SIGN SHAN COUNCIL TONE-4 (in the example to the left).

ဆမဲင်္ဂတအာင်း

ရွေးတဲမ် -အိုက်လွယ်ကန်

ခည်တလည်းခြိုဝ်းမျိုးအဲတအာင်း အဲဂေါင်းဒေး;ပေါ် ချိုင်ကဒည်အာရှတိုက်
ပေါ်နောဂေါင်းဇား မွန်ဂို အဲမိုးအဲနာည် မွန်ဂိုလျိုက်အုပ်စု မွန်ဗမာ အဲဂီဟာည်,
ဒီပေန်အဲမ်းကတာည်းဒီပဲန်း တတ်တီ, ဆမဲအဲဟိုမ်းအဲဟာ, စီလဲပုံစာမ်းလား;
ပေါ်ကိုင်းခယ်လေး; ဟာဝ်းညှာပစာန်း အဲမ်း ယန်စီကျန် စီလေး; တယ်လား;
ကိုင်းတန်ဟောင်။ စလာဘိုးပဒီး ဟာဝ်းလား;ပစာန်းစန်းတိုက် ယူနန်ပြည်နယ်
ကိုင်းမုန်ပန်း၊ မုန်ရန်၊ ပေါက်ဆန်၊ လိုင်ချမ်း၊ ကျန်းခမ်း၊ တက်ဟုန်ကျို
ရိမ်းကိုင်းမန်စီ (ဗေင်လာ) အိုင်ဇူးဟိုနဇာင် လှိုင်စာင်ဟောင် ဂေါင်းစောတ်လေး;ဒို့
တောန်းဒီး။ ကလ်ကီးဒိုဝ်းမားဟည့်ရပ်, ကမား;ဟာဝ်း။

အနောည် မွန်ဂို အလွိုက်ခြိုဝ်းတအာင်းအဲ စလား;ပါ နောဂေါင်း
ဇားမွန်ဂိုစီလဲပုံဂေါင်းလား;ပေါ် အရှေ့တောင်အာရှ။ ဓါည်းဒီးအဲစီဟာည်,
ဘိုးဒီပေန်ပဲန်းတတ်အဲဟိုမ်း ဇိုင်ကည့်အိုဝ် ဟာည်းဒီဇာကျာက်တမာည်းအိုဝ် ဒဲဝ်
ဆိုက်ဒဲဝ်နာန်း စီလောင်တော, လေး;လှုံအဲမဲတမာဝ်း ကာည်းအဲမ်း သံလွင်
ကဲအီလောင်လေး;တော, အဲမဲတမာဝ်း လဲပုံဟာဝ်းခဲင်းကိုင်းအိုင်ချာန်း
ဂေါင်းလေး;လား; ပေါ်ဟောင်ကိုင်းချေမ်း လှိုင်နမ့်ခမ်း၊ မူဆယ်၊ ကိုတိုခါည်၊
လားဆော်၊ နမ့်မတူ၊ မာန်တုံ၊ နမ့်ဆန်၊ ဆီပေါ၊ ကျောက်မဲ၊ မိုင်းကုတ်၊ မိုင်းမေတုံ
ဂေါင်းဟာဝ်း ကံပံတော်အိုင်ဒီး။

အီလောင်တော,လေး;အဲမ်း သံလွင် ကဲဒိုဝ်း စီဂေါင်းဟာဝ်းညှာ
စန်းဇောက် ကိုင်းချေမ်း လှိုင်မိုင်းယာင်း၊ မိုင်းခါက်၊ ကျင့်တုန်၊ မိုင်းကိုင်း၊ လဲချား၊
စေးတီး၊ မာန်တမ်၊ မိုင်းနောင်း၊ မိုးနဲ၊ ရပ်စောပ်၊ ကလော၊ အောင်ပါန်း၊ ပီင့်တိယ
ဂေါင်းဟာဝ်းကံပံတော်အိုင်ဒီးနီး။

အနောည်ကတူရ် အမျိုး ဆာမ်လုံ ကဲစီဒူည်းလေး; ထွင်းမဲမ်းတအာင်း
တော,လေး;လား;ချောင်း အဲမဲတမာဝ်း စီလေး;တယ်လား; ကိုင်းပေါ် နမ့်ဆန်
ဂေါင်းအူန်း ကံပံတော်အိုင်ဒီးနီး။

Figure 5. Sample from a Rumai Palaung reader published in 2005, showing ◉ LETTER RUMAI PALAUNG FA alongside ◌: MYANMAR SIGN DOT BELOW (Rumai Palaung tone 1), ◌: MYANMAR SIGN VISARGA (Rumai Palaung tone 3), ◌, SIGN SHAN TONE-2 (Rumai Palaung tone 4), ◌; SIGN SHAN TONE-3 (Rumai Palaung tone 5), and ◌; SIGN RUMAI PALAUNG TONE-6.

TABLE XX - Row 10: MYANMAR

	100	101	102	103	104	105	106	107	108	109
0	က	တ	ဣ	့	ဝ	ဝ	့		င့	
1	ခ	ထ	အ	ေ့	ဝ	ဝ	ှ		ံ	
2	ဂ	ဒ	က	ဲ	၂	ဗ	ါ		ံ	
3	ဃ	စ	ဒိ	ံ	၃	ဗ	ာ		ံ,	
4	င	န	ြိ	ဲ	၄	ဇ	ါ		ံ;	
5	စ	င	ဣ	ံ	၅	ဇ		ဂ	ံ°	
6	ဆ	ဇ	ဣ	ံ	၆	့		ဖ	ံ	
7	ဇ	ဇ	ဇ	ံ	၇	့		ဂ	ံ	
8	ဈ	တ	ဣ	ံ	၈	့		ဈ	ံ	
9	ည	ဖ	ြိ	ံ	၉	့		ဆ	ံ	
A	ဉ	ဃ	ြိ	ံ	၁	့		ဃ	ံ	
B	ဣ	ရ	ါ	ံ	၂	ဣ		ဣ		
C	ဣ	လ	ံ	ံ	ဣ	မ		ဃ		
D	ဣ	ဝ	ံ	ံ	ြိ	ဣ		၂		
E	ဃ	ဃ	ံ	ံ	၄	ံ		ံ		
F	က	တ	့	ယ	ဣ	ံ		ါ		

G = 00
P = 00

TABLE XX - Row 10: MYANMAR

hex	Name	hex	Name
00	MYANMAR LETTER KA	59	MYANMAR VOWEL SIGN VOCALIC LL
01	MYANMAR LETTER KHA	5A	MYANMAR LETTER MON NGA
02	MYANMAR LETTER GA	5B	MYANMAR LETTER MON JHA
03	MYANMAR LETTER GHA	5C	MYANMAR LETTER MON BBA
04	MYANMAR LETTER NGA	5D	MYANMAR LETTER MON BBE
05	MYANMAR LETTER CA	5E	MYANMAR CONSONANT SIGN MON MEDIAL NA
06	MYANMAR LETTER CHA	5F	MYANMAR CONSONANT SIGN MON MEDIAL MA
07	MYANMAR LETTER JA	60	MYANMAR CONSONANT SIGN MON MEDIAL LA
08	MYANMAR LETTER JHA	61	MYANMAR LETTER SGAW KAREN SHA
09	MYANMAR LETTER NYA	62	MYANMAR LETTER SGAW KAREN EU
0A	MYANMAR LETTER NNYA	63	MYANMAR SIGN SGAW KAREN HATHI
0B	MYANMAR LETTER TTA	64	MYANMAR SIGN SGAW KAREN KE PHO
0C	MYANMAR LETTER TTHA	65	(This position shall not be used)
0D	MYANMAR LETTER DDA	66	(This position shall not be used)
0E	MYANMAR LETTER DDHA	67	(This position shall not be used)
0F	MYANMAR LETTER NNA	68	(This position shall not be used)
10	MYANMAR LETTER TA	69	(This position shall not be used)
11	MYANMAR LETTER THA	6A	(This position shall not be used)
12	MYANMAR LETTER DA	6B	(This position shall not be used)
13	MYANMAR LETTER DHA	6C	(This position shall not be used)
14	MYANMAR LETTER NA	6D	(This position shall not be used)
15	MYANMAR LETTER PA	6E	(This position shall not be used)
16	MYANMAR LETTER PHA	6F	(This position shall not be used)
17	MYANMAR LETTER BA	70	(This position shall not be used)
18	MYANMAR LETTER BHA	71	(This position shall not be used)
19	MYANMAR LETTER MA	72	(This position shall not be used)
1A	MYANMAR LETTER YA	73	(This position shall not be used)
1B	MYANMAR LETTER RA	74	(This position shall not be used)
1C	MYANMAR LETTER LA	75	MYANMAR LETTER SHAN KA
1D	MYANMAR LETTER WA	76	MYANMAR LETTER SHAN KHA
1E	MYANMAR LETTER SA	77	MYANMAR LETTER SHAN CA
1F	MYANMAR LETTER HA	78	MYANMAR LETTER SHAN NYA
20	MYANMAR LETTER LLA	79	MYANMAR LETTER SHAN NA
21	MYANMAR LETTER A	7A	MYANMAR LETTER SHAN PHA
22	MYANMAR LETTER SHAN A	7B	MYANMAR LETTER SHAN FA
23	MYANMAR LETTER I	7C	MYANMAR LETTER SHAN THA
24	MYANMAR LETTER II	7D	MYANMAR LETTER SHAN HA
25	MYANMAR LETTER U	7E	MYANMAR CONSONANT SIGN SHAN MEDIAL WA
26	MYANMAR LETTER UU	7F	MYANMAR VOWEL SIGN SHAN AA
27	MYANMAR LETTER E	80	MYANMAR VOWEL SIGN SHAN E
28	MYANMAR LETTER MON E	81	MYANMAR VOWEL SIGN SHAN E ABOVE
29	MYANMAR LETTER O	82	MYANMAR VOWEL SIGN SHAN FINAL Y
2A	MYANMAR LETTER AU	83	MYANMAR SIGN SHAN TONE-2
2B	MYANMAR VOWEL SIGN TALL AA	84	MYANMAR SIGN SHAN TONE-3
2C	MYANMAR VOWEL SIGN AA	85	MYANMAR SIGN SHAN COUNCIL TONE-4
2D	MYANMAR VOWEL SIGN I	86	MYANMAR SIGN SHAN TONE-5
2E	MYANMAR VOWEL SIGN II	87	MYANMAR SIGN SHAN TONE-6
2F	MYANMAR VOWEL SIGN U	88	MYANMAR SIGN SHAN COUNCIL EMPHATIC TONE
30	MYANMAR VOWEL SIGN UU	89	MYANMAR VOWEL SIGN RUMAI PALAUNG FA
31	MYANMAR VOWEL SIGN E	8A	MYANMAR SIGN RUMAI PALAUNG TONE-5
32	MYANMAR VOWEL SIGN AI	8B	(This position shall not be used)
33	MYANMAR VOWEL SIGN MON II	8C	(This position shall not be used)
34	MYANMAR VOWEL SIGN MON O	8D	(This position shall not be used)
35	MYANMAR VOWEL SIGN E ABOVE	8E	(This position shall not be used)
36	MYANMAR SIGN ANUSVARA	8F	(This position shall not be used)
37	MYANMAR SIGN DOT BELOW	90	(This position shall not be used)
38	MYANMAR SIGN VISARGA	91	(This position shall not be used)
39	MYANMAR SIGN VIRAMA	92	(This position shall not be used)
3A	MYANMAR SIGN ASAT	93	(This position shall not be used)
3B	MYANMAR CONSONANT SIGN MEDIAL YA	94	(This position shall not be used)
3C	MYANMAR CONSONANT SIGN MEDIAL RA	95	(This position shall not be used)
3D	MYANMAR CONSONANT SIGN MEDIAL WA	96	(This position shall not be used)
3E	MYANMAR CONSONANT SIGN MEDIAL HA	97	(This position shall not be used)
3F	MYANMAR LETTER GREAT SA	98	(This position shall not be used)
40	MYANMAR DIGIT ZERO	99	(This position shall not be used)
41	MYANMAR DIGIT ONE	9A	(This position shall not be used)
42	MYANMAR DIGIT TWO	9B	(This position shall not be used)
43	MYANMAR DIGIT THREE	9C	(This position shall not be used)
44	MYANMAR DIGIT FOUR	9D	(This position shall not be used)
45	MYANMAR DIGIT FIVE	9E	(This position shall not be used)
46	MYANMAR DIGIT SIX	9F	(This position shall not be used)
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		

A. Administrative

1. Title

Proposal for encoding Myanmar characters for Shan and Palaung in the UCS.

2. Requester's name

UC Berkeley Script Encoding Initiative (Universal Scripts Project); authors: Michael Everson and Martin Hosken

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

Liaison contribution.

4. Submission date

2006-09-08

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

23

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, similar characters have been submitted before. See N2768, N3043, N3044, N3080

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?

Yes.

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

Some of the Shan tone marks look superficially like punctuation marks, but they are combining characters and typically have a hollow dot.

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?