

**Title:** Proposal of Myanmar Script Extensions:  
**Mon, Shan, and Karen (Kayin)**  
**Source:** Myanmar Unicode and NLP Research Center  
**Status:** National Contribution  
**Action:** For consideration by WG2  
**Date:** 19-May-2004

This document requests additional characters to be added to the UCS and contains the proposal summary form.

## A. Administrative

### 1. Title

Proposal of Myanmar Script Extensions: Mon, Shan, and Karen (Kayin)

### 2. Requester's name

Myanmar Unicode and NLP Research Center, Myanmar Computer Federation

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

National contribution.

### 4. Submission date

2004-05-19

### 5. Requester's reference (if applicable)

Thein OO, President, MCF <[mcf@mail4u.com.mm](mailto:mcf@mail4u.com.mm)>

Thein HTUT, Secretary, MCSA

Tun TINT, Member, Myanmar Language Commission

Zaw HTUT, Program Manager, Myanmar UNLP Research Center <[zhtut@myanmars.net](mailto:zhtut@myanmars.net)>

Ngwe TUN, Program Manager, Myanmar UNLP Research Center <[ngwestar@etrademyanmar.com](mailto:ngwestar@etrademyanmar.com)>

c/o:

#### **MYANMAR UNICODE AND NLP RESEARCH CENTER**

Myanmar Computer Federation

B1R1, Myanmar ICT Park, Universities' Hlaing Campus,

11052, Yangon, Myanmar

Tel: +95-1-652307

eFax: +1-707-988-0300

Email: [myanmar-nlp@mail4u.com.mm](mailto:myanmar-nlp@mail4u.com.mm) , [mcf@mail4u.com.mm](mailto:mcf@mail4u.com.mm)

Internet: <http://www.mcf.org.mm> , <http://myanmars.net/unicode/>

### 6. Choose one of the following:

#### 6a. This is a complete proposal

Yes.

#### 6b. More information will be provided later

Yes.

## 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.

#### 1b. Name of the existing block

Myanmar (U+1000 to U+109F)

**2. Number of characters in proposal**

30

**3. Proposed category (see section II, Character Categories)**

Category A.

**4a. Proposed Level of Implementation (1, 2 or 3) (see clause 14, ISO/IEC 10646-1: 2000)**

Level 3.

**4b. Is a rationale provided for the choice?**

Yes.

**4c. If YES, reference**

- Existing Mon, Shan, and Karen Primers
- Documents on <http://www.mcf.org.mm> , <http://myanmars.net/unicode/>

**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 ISO/IEC 10646-1: 2000?**

Not confirmed.

**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?**

Myanmar Unicode and NLP Research Center  
Zaw HTUT  
TrueType.

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

Zaw HTUT, Program Manager, Myanmar Unicode & NLP Research Center. Myanmar Computer Federation.  
email: [myanmar-nlp@mail4u.com.mm](mailto:myanmar-nlp@mail4u.com.mm) , [mcf@mail4u.com.mm](mailto:mcf@mail4u.com.mm)  
internet: <http://www.mcf.org.mm> , <http://myanmars.net/unicode/>  
FontLab.

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

Yes, see bibliography.

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

Yes.

**8. 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, the basics only.

**9. 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/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.**

Yes, see proposal below.

## C. Technical – Justification

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

No.

**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?**

- Myanmar ICT Standardization Steering Committee (national body)
- Mon, Shan, and Karen linguists who are working with the Myanmar Language Commission (national body)
- Myanmar Computer Federation [<http://www.mcf.org.mm>]
- Myanmar Computer Scientist Association [<http://www.mcsa.org.mm>]
- Myanmar Unicode and NLP Research Center [<http://myanmars.net/unicode>]

**2c. If YES, available relevant documents**

All available documents related to Myanmar NLP are listed at  
<http://www.mcf.org.mm> , <http://myanmars.net/unicode/doc/>

**3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?**

Yes.

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

Common everyday use.

**4b. Reference**

See bibliography.

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

Yes.

**5b. If YES, where?**

In Myanmar (formerly Burma), especially in Mon, Shan, and Karen States.

**6a. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing document) must the proposed characters be entirely in the BMP?**

Yes, since there is a reserved space for these.

**6b. If YES, is a rationale provided?**

Yes.

**6c. If YES, reference**

All Myanmar points are in the BMP.

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

Strongly suggest to be kept together with Myanmar.

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

Yet to be confirmed later.

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

Not yet.

**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?**

Still yet to be confirmed soon.

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

Not yet.

**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, some can be.

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

Yes.

**10c. If YES, reference**

Many of them derived from Pali and Sanskrit, but they have different functions and shapes.

**11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?**

Yes.

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

Yes.

**11c. If YES, reference**

Sample printed documents.

**12a. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?**

Yes.

**12b. If YES, reference**

See Annex-3.

**13a. Does the proposal contain characters with any special properties such as control function or similar semantics?**

No.

**13b. If YES, describe in detail (include attachment if necessary)**

**14a. Does the proposal contain any Ideographic compatibility character(s)?**

No.

**14b. If YES, is the equivalent corresponding unified ideographic character(s) identified?**

**14c. If YES, reference**

## D. Proposal

### D.1. Background

D.1.1. Myanmar script was firstly proposed to Unicode Consortium in 1992, and was successfully encoded in the repertoire in 1998 in Unicode Standard 3.0. Other sub-scripts, such as Kayah, Karen, Chin, Mon, and Shan, based on the Myanmar script.

D.1.2. Out of 160 code points reserved for the Myanmar script (between U+1000 and U+109F), only 78 code points were encoded and used for Myanmar script, as of the Unicode 4.0.

D.1.3. Totally 82 code points in the Myanmar block in the Unicode BMP (Basic Multilingual Plane) are still being reserved for the rest of the subscripts to be encoded.

### D.2. User Community

D.2.1. The Mon state is inhabited by approximately 2,000,000 people, and its size is 4748 square miles. A lot of Mons still live in western Thailand, near Myanmar border. The use of IT is moderate, and the number of Mon books published a year is not available.

D.2.2. The Shan state is the largest state in the country, and the estimated number of inhabitants is as large as 4,400,000. The size of Shan state is 60155 square miles. Due to the nature of high Shan plateau, most parts are remote areas, and IT usage is very limited. Shan culture is closer to the Siamese, now called Thai. A lot of Shan people also live in Thailand and China. Publications printed in Shan script may be the largest after Myanmar. Its demographics shows that the Shans have 33 minor ethnic groups.

D.2.3. The Karen state is bordering Thailand and has 1,200,000 inhabitants approximately. The size 11,731 square miles is not the smallest state in Myanmar. Use of IT in schools and offices are quite limited. It was noted that the Christian bible was even translated and then printed in Karen language. Regarding demographics, there are 11 minor ethnic groups in Karen group, and among them Sakaw Karen and Poe Karen are dominant and their scripts have a few minor differences.

### D.3. Proposed Characters

#### D.3.1. Mon Sub-Script

D.3.1.1. Actually it was the Mons who brought in the Indict script first and applied it for their own language. Myanmar follows later. Mons use all the alphabets of the Myanmar, plus 2 more consonants.

D.3.1.2. Number of new proposed Mon characters: 6 (3 consonants, 2 medials (combining consonants), and 1 special character)

D.3.1.3. Graphical presentation of proposed Mon character set: See [Annex-1](#).

D.3.1.4. List of proposed Mon character names: See [Annex-2](#).

#### D.3.2. Shan Sub-Script

D.3.2.1. Shan script seems to be a lot more simplified than Myanmar and Thai, having only 20 consonants.

D.3.2.2. Number of new proposed Shan characters: 14 (2 medials (combining consonants), and 5 vowel sign and 7 tone mark sign)

D.3.2.3. Graphical presentation of proposed Shan character set: See [Annex-3](#).

D.3.2.4. List of proposed Shan character names: See [Annex-4](#).

#### D.3.3. Karen Sub-Script

D.3.3.1. This presented version of Karen script is a combined set of characters both from Sakaw Karen and Poe Karen scripts.

D.3.3.2. Number of new proposed Karen characters: 10 (2 consonants, 2 medials (combining consonants), and 6 devowelisers)

D.3.3.3. Graphical presentation of proposed Karen character set: See [Annex-5](#).

D.3.3.4. List of proposed Karen character names: See [Annex-6](#).

#### **D.4. Character Properties of Proposed Characters**

##### D.4.1. Line-breaking Behavior

The line-breaking rules applies same as the base script -- Myanmar. It is true that Myanmar script does not use "white space" like Latin to break up words. But Myanmar line breaking is not as complicated as Thai.

According to Myanmar script, syllables have be formed up, and all line breaks take place ONLY outside syllable boundaries (always before or after syllables). To clarify, the line-breaks takes place always before pre-base and base characters, and always after a set of post-base characters.

It is important to note that there are pre-base characters before the base consonants, although they have be stored behind the base character in the backing store.

##### D.4.2. Combing Behavior and Spacing Behaviors

D.4.2.1. All consonants and special vowel characters have spacing same as any other consonants.

D.4.2.2. Above-base, below-base, and stacked characters (as modifiers of the base character) does not take up any spacing.

D.4.2.3. Some post-base characters takes up a little bit of spacing.

#### D.4.2.4. Mon characters

Mon: (base)



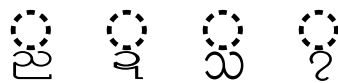
Mon: (above-base)



Mon: (below-base)



Mon: (stacked)



#### D.4.2.5. Mon Consonants



1004



1008



1060



1061

#### D.4.2.6 Mon Medial



1063



1064



1065

#### D.4.2.7. Mon Vowel Signs



1024



1027



1062

**D.4.2.8. Shan characters**

Shan: (base)



Shan: (pre-base)



Shan: (post-base)



Shan: (below-base)



**D.4.2.9. Shan Consonants**

ၵ	ၶ	ၷ	ၸ	ၹ
1000	1001	1005	1006	100A
ၶ̃	ၷ̃	ၸ̃	ၹ̃	
1014	1016	101F	1021	

**D.4.2.10. Shan Medial Characters**

ၵ̅	ၶ̅
1068	1069

**D.4.2.11. Shan Vowel Signs**

ၵ̃̄	ၶ̃̄	ၷ̃̄	ၸ̃̄	ၹ̃̄
106A	106B	106C	106D	106E





**D.4.2.12. Shan Tonal Marks**


ၵ̄̄	ၵ̄̄°	ၵ̄̄°°	ၵ̄̄°	ၵ̄̄°°
106F	1070-	1071	1072	1073


**D.4.2.13. Shan Paoh Tonal Marks**

ၵ̄̄̄	ၵ̄̄̄
1066	1067


#### D.4.2.14. Karen characters

Karen: (base)                        

Karen: (post-base)      

Karen: (below-base)      

#### D.4.2.15. Karen Consonants




1074      1075

#### D.4.2.16. Karen Medial Characters




      

1065      1076

#### D.4.2.17. Karen Devowelisers (killers)

1077      1078      1079

107A      107B      107C

#### D.4.4. Directional Behavior

All Myanmar scripts and sub-scripts are written from left to right, then top to bottom, exactly as Latin does. Needless to say that the numerals are written from right to left, as same as Arabic numerals used in Latin encoding.

#### D.4.5. Default Collation Behavior:

n.a.

D.4.6. There is nothing or no available data at the moment to submit on casing, numeric, default collation behaviors, relevance at markup context, compatibility equivalence, and other Unicode normalization.

D.4.7. Currency information is submitted in a separate proposal.

## E. Bibliography

1. Myanmar-Myanmar-English Dictionary  
by Myanmar Language Commission
2. Myanmar Orthography  
by Myanmar Language Commission, 1999 Edition
3. Mon-English Dictionary  
by Halliday
4. English-Mon Dictionary  
by Reverend Stevens, ....
5. Shan-English Dictionary  
by Cushing
6. The Unicode Standard 3.0
7. The Unicode Standard 4.0  
by Unicode Consortium, 2002
8. Karen Primer
9. Mon Primer
10. Shan Primer
11. A Comparative Study of Shan, Poe Karen, Sakaw Karen, and Paoh Scripts  
by
12. Collection of Myanmar NLP Research Work and Papers  
<http://myanmars.net/unicode/>

**ANNEX-1**

1000

Myanmar(Mon)

109F

	100	101	102	103	104	105	106	107	108	109
0	က 1000	တ 1010	င 1020	ူ 1030	ဝ 1040	ဂ 1050	ဇ 1060			
1	ခ 1001	ထ 1011	အ 1021	ေ 1031	၁ 1041	ဗ 1051	မှ 1061			
2	ဂ 1002	ဒ 1012		ဲ 1002	၂ 1042	ဖ 1052	် 1062			
3	ဃ 1003	ဓ 1013	ဒိ 1023		၃ 1043	ဗ 1053	န 1063			
4	င 1004	န 1014	ဒိ 1024		၄ 1044	ဇ 1054	န 1064			
5	စ 1005	ပ 1015	ည 1025		၅ 1045	ဇ 1055	၇ 1065			
6	ဆ 1006	ဖ 1016	ည 1026	ံ 1036	၆ 1046	ယ 1056				
7	ဇ 1007	ဗ 1017	ဇ 1027	့ 1037	၇ 1047	ယ 1057				
8	ဇ 1008	ဘ 1018		း 1038	၈ 1048	ဇ 1058				
9	ည 1009	မ 1019	သ 1029	့ 1039	၉ 1049	ဇ 1059				
A	ည 100A	ယ 101A	သ 102A		I 104A					
B	င 100B	ရ 101B			II 104B					
C	င 100C	လ 101C	ာ 102C		၂ 104C					
D	ည 100D	ဝ 101D	ံ 102D		၃ 104D					
E	ပ 100E	သ 101E	ံ 102E		၄ 104E					
F	က 100F	တ 101F	ူ 102F		၅ 104F					

**ANNEX-2**

**1000**

**Consonants**

1000	က	MYANMAR LETTER KA
1001	ခ	MYANMAR LETTER KHA
1002	ဂ	MYANMAR LETTER GA
1003	ဃ	MYANMAR LETTER GHA
1004	င	MYANMAR LETTER NGA
1005	စ	MYANMAR LETTER CA
1006	ဆ	MYANMAR LETTER CHA
1007	ဇ	MYANMAR LETTER JA
1008	ဇာ	MYANMAR LETTER JHA
1009	ည	MYANMAR LETTER NYA
100A	ညာ	MYANMAR LETTER NNYA
100B	တ	MYANMAR LETTER TTA
100C	တာ	MYANMAR LETTER TTHA
100D	ဒ	MYANMAR LETTER DDA
100E	ဗ	MYANMAR LETTER DDHA
100F	ဏ	MYANMAR LETTER NNA
1010	တ	MYANMAR LETTER TA
1011	ထ	MYANMAR LETTER THA
1012	ဒ	MYANMAR LETTER DA
1013	ဓ	MYANMAR LETTER DHA
1014	န	MYANMAR LETTER NA
1015	ပ	MYANMAR LETTER PA
1016	ဖ	MYANMAR LETTER PHA
1017	ဗ	MYANMAR LETTER BHA
1018	ဘ	MYANMAR LETTER BA
1019	မ	MYANMAR LETTER MA
101A	ယ	MYANMAR LETTER YA
101B	ရ	MYANMAR LETTER RA
101C	လ	MYANMAR LETTER LA
101D	ဝ	MYANMAR LETTER WA
101E	သ	MYANMAR LETTER SA
101F	ဟ	MYANMAR LETTER HA
1020	ဠ	MYANMAR LETTER LLA

**Independent vowels**

1021	အ	MYANMAR LETTER A • also represents the glottal stop as a consonant
1022		<RESERVED>
1023	အိ	MYANMAR LETTER I
1024	အီ	MYANMAR LETTER MON II
1025	ဣ	MYANMAR LETTER U
1026	ဤ	MYANMAR LETTER UU ≡1025 ဥ 102E ့
1027	ဧ	MYANMAR LETTER MON E
1028		<RESERVED>
1029	အို	MYANMAR LETTER O
102A	အို	MYANMAR LETTER AU

**Dependent vowels**

102C	ာ	MYANMAR VOWEL SIGN AA
102D	ိ	MYANMAR VOWEL SIGN I

**Myanmar(Mon)**

102E	ီ	MYANMAR VOWEL SIGN II
102F	ု	MYANMAR VOWEL SIGN U
1030	ူ	MYANMAR VOWEL SIGN UU
1031	ေ	MYANMAR VOWEL SIGN E • stands to the left of the consonant
1032	ဲ	MYANMAR VOWEL SIGN AI

**Various signs**

1036	ံ	MYANMAR SIGN ANUSVARA
1037	့	MYANMAR SIGN DOT BELLOW =aukmyit • a tone mark
1038	ံ့	MYANMAR SIGN VISARGA
1039	်	MYANMAR SIGN VIRAMA =killer (when rendered visibly)

**Digits**

1040	၀	MYANMAR DIGIT ONE
1041	၁	MYANMAR DIGIT TWO
1042	၂	MYANMAR DIGIT THREE
1043	၃	MYANMAR DIGIT FOUR
1044	၄	MYANMAR DIGIT FIVE
1045	၅	MYANMAR DIGIT SIX
1046	၆	MYANMAR DIGIT SEVEN
1047	၇	MYANMAR DIGIT EIGHT
1048	၈	MYANMAR DIGIT NINE
1049	၉	MYANMAR DIGIT TEN

**Punctuation**

104A	၊	MYANMAR SIGN LITTLE SECTION → 0964 ၊ devanagari danda
104B	။	MYANMAR SIGN SECTION → 0965 ။ devanagari double danda

**Various signs**

104C	ံ့	MYANMAR SYMBOL LOCATIVE
104D	ံ့	MYANMAR SYMBOL VOMPLETED
104E	ံ့	MYANMAR SYMBOL AMFOREMETION
104F	ံ့	MYANMAR SYMBOL GENITIVE

**Pali and Sanskrit extensions**

1050	ခ	MYANMAR LETTER SHA
1051	ဗ	MYANMAR LETTER SSA
1052	ဗ	MYANMAR LETTER VOCALIC R
1053	ဗ	MYANMAR LETTER VOCALIC RR
1054	ဗ	MYANMAR LETTER VOCALIC L
1055	ဗ	MYANMAR LETTER VOCALIC LL
1056	ံ	MYANMAR VOWEL SIGN VOCALIC R
1057	ံ	MYANMAR VOWEL SIGN VOCALIC RR
1058	ံ	MYANMAR VOWEL SIGN VOCALIC L
1059	ံ	MYANMAR VOWEL SIGN VOCALIC LL

**Mon Extension**

1060	ဝ	MON LETTER BA.
1061	မ	MON LETTER BE.
1062	ံ	MON VOWEL AO
1063	ံ	MON MEDIAL NA
1064	ံ	MON MEDIAL MA
1065	ံ	MON MEDIAL LA

**ANNEX-3**

1000

Myanmar(Shan)

109F

	100	101	102	103	104	105	106	107	108	109
0	၈ 1000	၉ 1010	၁၀ 1020	၁၁ 1030	၁၂ 1040	၁၃ 1050		၁၄ 1070		
1	၁၅ 1001	၁၆ 1011	၁၇ 1021	၁၈ 1031	၁၉ 1041	၂၀ 1051		၂၁ 1071		
2	၂၂ 1002	၂၃ 1012		၂၄ 1032	၂၅ 1042	၂၆ 1052		၂၇ 1072		
3	၂၈ 1003	၂၉ 1013	၃၀ 1023		၃၁ 1043	၃၂ 1053		၃၃ 1073		
4	၃၄ 1004	၃၅ 1014	၃၆ 1024		၃၇ 1044	၃၈ 1054				
5	၃၉ 1005	၄၀ 1015	၄၁ 1025		၄၂ 1045	၄၃ 1055				
6	၄၄ 1006	၄၅ 1016	၄၆ 1026	၄၇ 1036	၄၈ 1046	၄၉ 1056	၅၀ 1066			
7	၅၁ 1007	၅၂ 1017	၅၃ 1027	၅၄ 1037	၅၅ 1047	၅၆ 1057	၅၇ 1067			
8	၅၈ 1008	၅၉ 1018		၆၀ 1038	၆၁ 1048	၆၂ 1058	၆၃ 1068			
9	၆၄ 1009	၆၅ 1019	၆၆ 1029	၆၇ 1039	၆၈ 1049	၆၉ 1059	၇၀ 1069			
A	၇၁ 100A	၇၂ 101A	၇၃ 102A		၇၄ 104A		၇၅ 106A			
B	၇၆ 100B	၇၇ 101B			၇၈ 104B		၇၉ 106B			
C	၈၀ 100C	၈၁ 101C	၈၂ 102C		၈၃ 104C		၈၄ 106C			
D	၈၅ 100D	၈၆ 101D	၈၇ 102D		၈၈ 104D		၈၉ 106D			
E	၉၀ 100E	၉၁ 101E	၉၂ 102E		၉၃ 104E		၉၄ 106E			
F	၉၅ 100F	၉၆ 101F	၉၇ 102F		၉၈ 104F		၉၉ 106F			

**ANNEX-3**

**1000**

**Consonants**

1000	ခ	MYANMAR LETTER KA
1001	စ	MYANMAR LETTER KHA
1002	ဂ	MYANMAR LETTER GA
1003	ဃ	MYANMAR LETTER GHA
1004	င	MYANMAR LETTER NGA
1005	စ	MYANMAR LETTER CA
1006	ဆ	MYANMAR LETTER CHA
1007	ဇ	MYANMAR LETTER JA
1008	ဈ	MYANMAR LETTER JHA
1009	ည	MYANMAR LETTER NYA
100A	ည	MYANMAR LETTER NNYA
100B	တ	MYANMAR LETTER TTA
100C	တ	MYANMAR LETTER TTHA
100D	ဒ	MYANMAR LETTER DDA
100E	ဒ	MYANMAR LETTER DDHA
100F	န	MYANMAR LETTER NNA
1010	တ	MYANMAR LETTER TA
1011	ထ	MYANMAR LETTER THA
1012	ဒ	MYANMAR LETTER DA
1013	ဓ	MYANMAR LETTER DHA
1014	န	MYANMAR LETTER NA
1015	ပ	MYANMAR LETTER PA
1016	ဖ	MYANMAR LETTER PHA
1017	ပ	MYANMAR LETTER BHA
1018	ဘ	MYANMAR LETTER BA
1019	မ	MYANMAR LETTER MA
101A	ယ	MYANMAR LETTER YA
101B	ရ	MYANMAR LETTER RA
101C	လ	MYANMAR LETTER LA
101D	ဝ	MYANMAR LETTER WA
101E	သ	MYANMAR LETTER SA
101F	ဟ	MYANMAR LETTER HA
1020	ဗ	MYANMAR LETTER LLA

**Independent vowels**

1021	အ	MYANMAR LETTER A • also represents the glottal stop as a consonant
1022		<RESERVED>
1023	ဣ	MYANMAR LETTER I
1024	ဩ	MYANMAR LETTER MON II
1025	ဪ	MYANMAR LETTER U
1026	ါ	MYANMAR LETTER UU ≡1025 ဩ 102E ဩ
1027	ဥ	MYANMAR LETTER MON E
1028		<RESERVED>
1029	ဪ	MYANMAR LETTER O
102A	ဩ	MYANMAR LETTER AU

**Dependent vowels**

102C	◌◌	MYANMAR VOWEL SIGN AA
102D	◌◌	MYANMAR VOWEL SIGN I
102E	◌◌	MYANMAR VOWEL SIGN II
102F	◌◌	MYANMAR VOWEL SIGN U

**Myanmar(Shan)**

1030	◌◌	MYANMAR VOWEL SIGN UU
1031	◌◌	MYANMAR VOWEL SIGN E • stands to the left of the consonant
1032	◌◌	MYANMAR VOWEL SIGN AI

**Various signs**

1036	◌◌	MYANMAR SIGN ANUSVARA
1037	◌◌	MYANMAR SIGN DOT BELLOW =aukmyit • a tone mark
1038	◌◌	MYANMAR SIGN VISARGA
1039	◌◌	MYANMAR SIGN VIRAMA =killer (when rendered visibly)

**Digits**

1040	၀	MYANMAR DIGIT ONE
1041	၁	MYANMAR DIGIT TWO
1042	၂	MYANMAR DIGIT THREE
1043	၃	MYANMAR DIGIT FOUR
1044	၄	MYANMAR DIGIT FIVE
1045	၅	MYANMAR DIGIT SIX
1046	၆	MYANMAR DIGIT SEVEN
1047	၇	MYANMAR DIGIT EIGHT
1048	၈	MYANMAR DIGIT NINE
1049	၉	MYANMAR DIGIT TEN

**Punctuation**

104A	၊	MYANMAR SIGN LITTLE SECTION → 0964 ၊ devanagari danda
104B	။	MYANMAR SIGN SECTION → 0965 ။ devanagari double danda

**Various signs**

104C	◌◌	MYANMAR SYMBOL LOCATIVE
104D	◌◌	MYANMAR SYMBOL VOMPLETED
104E	◌◌	MYANMAR SYMBOL AMFOREMETION
104F	◌◌	MYANMAR SYMBOL GENITIVE

**Pali and Sanskrit extensions**

1050	◌◌	MYANMAR LETTER SHA
1051	◌◌	MYANMAR LETTER SSA
1052	◌◌	MYANMAR LETTER VOCALIC R
1053	◌◌	MYANMAR LETTER VOCALIC RR
1054	◌◌	MYANMAR LETTER VOCALIC L
1055	◌◌	MYANMAR LETTER VOCALIC LL
1056	◌◌	MYANMAR VOWEL SIGN VOCALIC R
1057	◌◌	MYANMAR VOWEL SIGN VOCALIC RR
1058	◌◌	MYANMAR VOWEL SIGN VOCALIC L
1059	◌◌	MYANMAR VOWEL SIGN VOCALIC LL

**Shan Extension**

1066	◌◌	SHAN TONE MARK MAI PAK NGA
1067	◌◌	SHAN TONE MARK MAI NGA
1068	◌◌	SHAN MEDIAL VA
1069	◌◌	SHAN MEDIAL PHA
106A	◌◌	SHAN VOWEL SIGN AOU
106B	◌◌	SHAN VOWEL SIGN SUPERSCRRIPT A
106C	◌◌	SHAN VOWEL SIGN SUPERSCRRIPT AY
106D	◌◌	SHAN VOWEL SIGN SUPERSCRRIPT AI
106E	◌◌	SHAN VOWEL SIGN SUPERSCRRIPT AIM
106F	◌◌	SHAN TONE MARK AA
1070	◌◌	SHAN TONE MARK AI
1071	◌◌	SHAN TONE MARK AII
1072	◌◌	SHAN TONE MARK BELOW AI
1073	◌◌	SHAN TONE MARK BELOW AII

ANNEX-5

1000 Myanmar(Karen) 109F

	100	101	102	103	104	105	106	107	108	109
0	က 1000	တ 1010	င 1020	ု 1030	ဝ 1040	ဂ 1050				
1	ခ 1001	ဖ 1011	အ 1021	ေ 1031	၁ 1041	ဗ 1051				
2	ဂ 1002	ဒ 1012		ဲ 1002	၂ 1042	ဇ 1052				
3	ဃ 1003	စ 1013	ဒိ 1023		၃ 1043	ဗ 1053				
4	င 1004	န 1014	ဩ 1024		၄ 1044	ဇ 1054		၇ 1074		
5	စ 1005	ပ 1015	၇ 1025		၅ 1045	ဇ 1055	၈ 1065	၈ 1075		
6	ဆ 1006	ဖ 1016	၇ 1026	ံ 1036	၆ 1046	ယ 1056		၉ 1076		
7	ဇ 1007	ပ 1017	ဇ 1027	့ 1037	၇ 1047	ယ 1057		ါ 1077		
8	ဈ 1008	ဘ 1018		း 1038	၈ 1048	ဉ 1058		ာ 1078		
9	၇ 1009	မ 1019	သ 1029	့ 1038	၉ 1049	ဉ 1059		း 1079		
A	ည 100A	ယ 101A	သွ 102A		ါ 104A			ာ 107A		
B	ဏ 100B	ရ 101B			ါ 104B			ာ 107B		
C	င 100C	လ 101C	ာ 102C		ာ 104C			ာ 107C		
D	၇ 100D	ဝ 101D	ိ 102D		ာ 104D					
E	ပ 100E	သ 101E	ိ 102E		ာ 104E					
F	က 100F	တ 101F	ု 102F		ာ 104F					

**ANNEX-3**

**1000**

**Consonants**

1000	က	MYANMAR LETTER KA
1001	ခ	MYANMAR LETTER KHA
1002	ဂ	MYANMAR LETTER GA
1003	ဃ	MYANMAR LETTER GHA
1004	င	MYANMAR LETTER NGA
1005	စ	MYANMAR LETTER CA
1006	ဆ	MYANMAR LETTER CHA
1007	ဇ	MYANMAR LETTER JA
1008	ဈ	MYANMAR LETTER JHA
1009	ည	MYANMAR LETTER NYA
100A	ဋ	MYANMAR LETTER NNYA
100B	တ	MYANMAR LETTER TTA
100C	ထ	MYANMAR LETTER TTHA
100D	ဒ	MYANMAR LETTER DDA
100E	ဗ	MYANMAR LETTER DDHA
100F	ဏ	MYANMAR LETTER NNA
1010	တ	MYANMAR LETTER TA
1011	ထ	MYANMAR LETTER THA
1012	ဒ	MYANMAR LETTER DA
1013	ဓ	MYANMAR LETTER DHA
1014	န	MYANMAR LETTER NA
1015	ပ	MYANMAR LETTER PA
1016	ဖ	MYANMAR LETTER PHA
1017	ဗ	MYANMAR LETTER BHA
1018	ဘ	MYANMAR LETTER BA
1019	မ	MYANMAR LETTER MA
101A	ယ	MYANMAR LETTER YA
101B	ရ	MYANMAR LETTER RA
101C	လ	MYANMAR LETTER LA
101D	ဝ	MYANMAR LETTER WA
101E	သ	MYANMAR LETTER SA
101F	ဟ	MYANMAR LETTER HA
1020	ဠ	MYANMAR LETTER LLA

**Independent vowels**

1021	အ	MYANMAR LETTER A • also represents the glottal stop as a consonant
1022		<RESERVED>
1023	ဣ	MYANMAR LETTER I
1024	ဤ	MYANMAR LETTER MON II
1025	ဥ	MYANMAR LETTER U
1026	ဦ	MYANMAR LETTER UU ≡1025 ဥ 102E ဦ
1027	ဧ	MYANMAR LETTER MON E
1028		<RESERVED>
1029	ဩ	MYANMAR LETTER O
102A	ဪ	MYANMAR LETTER AU

**Dependent vowels**

102C	ာ	MYANMAR VOWEL SIGN AA
102D	ိ	MYANMAR VOWEL SIGN I
102E	ီ	MYANMAR VOWEL SIGN II
102F	ု	MYANMAR VOWEL SIGN U

**Myanmar(Karen)**

1030	ူ	MYANMAR VOWEL SIGN UU
1031	့	MYANMAR VOWEL SIGN E • stands to the left of the consonant
1032	ဲ	MYANMAR VOWEL SIGN AI

**Various signs**

1036	ံ	MYANMAR SIGN ANUSVARA
1037	့	MYANMAR SIGN DOT BELLOW =aukmyit • a tone mark
1038	း	MYANMAR SIGN VISARGA
1039	်	MYANMAR SIGN VIRAMA =killer (when rendered visibly)

**Digits**

1040	၀	MYANMAR DIGIT ONE
1041	၁	MYANMAR DIGIT TWO
1042	၂	MYANMAR DIGIT THREE
1043	၃	MYANMAR DIGIT FOUR
1044	၄	MYANMAR DIGIT FIVE
1045	၅	MYANMAR DIGIT SIX
1046	၆	MYANMAR DIGIT SEVEN
1047	၇	MYANMAR DIGIT EIGHT
1048	၈	MYANMAR DIGIT NINE
1049	၉	MYANMAR DIGIT TEN

**Punctuation**

104A	၊	MYANMAR SIGN LITTLE SECTION → 0964 ၊ devanagari danda
104B	။	MYANMAR SIGN SECTION → 0965 ။ devanagari double danda

**Various signs**

104C	း	MYANMAR SYMBOL LOCATIVE
104D	ှ်	MYANMAR SYMBOL VOMPLETED
104E	း	MYANMAR SYMBOL AMFOREMETION
104F	၏	MYANMAR SYMBOL GENITIVE

**Pali and Sanskrit extensions**

1050	၀	MYANMAR LETTER SHA
1051	၁	MYANMAR LETTER SSA
1052	၂	MYANMAR LETTER VOCALIC R
1053	၃	MYANMAR LETTER VOCALIC RR
1054	၄	MYANMAR LETTER VOCALIC L
1055	၅	MYANMAR LETTER VOCALIC LL
1056	ံ	MYANMAR VOWEL SIGN VOCALIC R
1057	ံ	MYANMAR VOWEL SIGN VOCALIC RR
1058	ံ	MYANMAR VOWEL SIGN VOCALIC L
1059	ံ	MYANMAR VOWEL SIGN VOCALIC LL

**Karen Extension**

1065	၀	KAYIN MEDIAL LA
1074	၁	KAYIN LETTER SHAA
1075	၂	KAYIN LETTER HAA
1076	၃	KAYIN MEDIAL YA
1077	ံ	KAYIN DEVOWELISER EI THII
1078	ံ	KAYIN DEVOWELISER A' THII
1079	ံ	KAYIN DEVOWELISER HPLE HSI
107A	ံ	KAYIN DEVOWELISER HA THII
107B	ံ	KAYIN DEVOWELISER KAY PHO
107C	ံ	KAYIN DEVOWELISER KHLA THII