

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

L2/04-198

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>

Thein HTUT, Secretary, MCSA

Tun TINT, Member, Myanmar Language Commission

Zaw HTUT, Program Manager, Myanmar UNLP Research Center <zhtut@myanmars.net>

Ngwe TUN, Program Manager, Myanmar UNLP Research Center <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 , 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.

email: myanmar-nlp@mail4u.com.mm, mcf@mail4u.com.mm,

internet: <http://www.mcf.org.mm>, <http://www.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 close 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 break 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) doesnot 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)

၂ ခ ဃ င

Shan: (pre-base)

၆

Shan: (post-base)

၇ ၈ ၉

Shan: (below-base)

၁၀ ၁၁ ၁၂ ၁၃

D.4.2.9. Shan Consonants

၂

1000

ခ

1001

ဃ

1005

င

1006

၆

100A

၇

1014

၈

1016

၉

101F

၁၀

1021

D.4.2.10. Shan Medial Characters

၁၁

1068

၁၂

1069

D.4.2.11. Shan Vowel Signs

၁၃

106A

၁၄

106B

၁၅

106C

၁၆

106D

၁၇

106E

D.4.2.12. Shan Tonal Marks

၁၈

106F

၁၉

1070-

၂၀

1071

၂၁

1072

၂၂

1073

D.4.2.13. Shan Paoh Tonal Marks



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/>

1000

Myanmar(Mon)

	100	101	102	103	104	105	106	107	108	109
0	က 1000	တ 1010	င 1020	ါ 1030	ဝ 1040	ဂ 1050	ဓ 1060			
1	ခ 1001	ထ 1011	အ 1021	ေ 1031	၁ 1041	ဖ 1051	မှ 1061			
2	ဂ 1002	ဒ 1012		ဲ 1002	၂ 1042	ဇ 1052	် 1062			
3	ဃ 1003	မ 1013	ဒိ 1023		၃ 1043	ဗ 1053	ွ 1063			
4	ငှ 1004	န 1014	ဒိ 1024		၄ 1044	ဇ 1054	ွ 1064			
5	စ 1005	ပ 1015	ဥ 1025		၅ 1045	ဇ 1055	ွ 1065			
6	ဆ 1006	ဖ 1016	ဥ 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	ွ 1058				
9	ဥ 1009	မ 1019	သြ 1029		၉ 1049	ွ 1059				
A	ဥ 100A	ယ 101A	သြ 102A		I 104A					
B	ဥ 100B	ရ 101B			II 104B					
C	ဥ 100C	လ 101C	ာ 102C		၂ 104C					
D	ဥ 100D	ဝ 101D	ိ 102D		၃ 104D					
E	ဥ 100E	သ 101E	ိ 102E		၄ 104E					
F	က 100F	ဟ 101F	ါ 102F		ပ 104F					

1000

Consonant

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	ည	MYANMAR LETTER JHA
1009	ဉ	MYANMAR LETTER NYA
100A	ည	MYANMAR LETTER NNYA
100B	တ	MYANMAR LETTER TTA
100C	ထ	MYANMAR LETTER TTHA
100D	ဒ	MYANMAR LETTER DDA
100E	ဍ	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	ဩ	MYANMAR LETTER I
1024	ဪ	MYANMAR LETTER MON II
1025	ါ	MYANMAR LETTER U
1026	ာ	MYANMAR LETTER UU ≡ 1025 ဥ 102E ့
1027	ဧ	MYANMAR LETTER MON E
1028		<RESERVED>
1029	ဪ	MYANMAR LETTER O
102A	ါ	MYANMAR LETTER AU
102B		<RESERVED>

Dependent vowel signs

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

Myanmar(Mon)

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 ၊ dev anagari danda
104B	။	MYANMAR SIGN SECTION → 0965 ။ dev anagari 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 LETTER SIGN VOCALIC R
1057	၎	MYANMAR LETTER SIGN VOCALIC RR
1058	၎	MYANMAR LETTER SIGN VOCALIC L
1059	၎	MYANMAR LETTER SIGN VOCALIC LL

Mon extensions

1060	၎	MON LETTER BA
1061	၎	MON LETTER BE
1062	၎	MON VOWEL SIGN AO
1063	၎	MON MEDIAL NA
1064	၎	MON MEDIAL MA
1065	၎	MON MEDI

1000

Myanmar(Shan)

	100	101	102	103	104	105	106	107	108	109
0	၈ 1000	တ 1010	င 1020	ူ 1030	ဝ 1040	ဂ 1050		ံ 1070		
1	ခ 1001	ထ 1011	က 1021	ေ 1031	၁ 1041	ဗ 1051		ံ 1071		
2	ဂ 1002	ဒ 1012		ဲ 1032	၂ 1042	ဖ 1052		ံ 1072		
3	ဃ 1003	စ 1013	ဒိ 1023		၃ 1043	ဗ 1053		ံ 1073		
4	င 1004	ဆ 1014	ညိ 1024		၄ 1044	ဇ 1054				
5	လ 1005	ပ 1015	ည 1025		၅ 1045	ဇ 1055				
6	ဆ 1006	ဃ 1016	ည 1026	ံ 1036	၆ 1046	ယ 1056	ံ 1066			
7	ဇ 1007	ပ 1017	ဇ 1027	ံ 1037	၇ 1047	ယ 1057	ံ 1067			
8	ဈ 1008	တ 1018		ံ 1038	၈ 1048	ွ 1058	ွ 1068			
9	ည 1009	မ 1019	သ 1029		၉ 1049	ွ 1059	ွ 1069			
A	လု 100A	ယ 101A	သိ 102A		I 104A		ေ 106A			
B	ဋ 100B	ရ 101B			II 104B		ံ 106B			
C	ငှ 100C	လ 101C	ာ 102C		၂ 104C		ံ 106C			
D	ည 100D	ဝ 101D	ံ 102D		၃ 104D		ံ 106D			
E	ပ 100E	သ 101E	ံ 102E		၄ 104E		ံ 106E			
F	က 100F	၇ 101F	ူ 102F		ပ 104F		ဲ 106F			

1000

Consonant

1000	က	MYANMAR LETTER KA
1001	ခ	MYANMAR LETTER KHA
1002	ဂ	MYANMAR LETTER GA
1003	ဃ	MYANMAR LETTER GHYA
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	ဍ	MYANMAR LETTER TTA
100C	ဎ	MYANMAR LETTER TTHA
100D	ဏ	MYANMAR LETTER DDA
100E	တ	MYANMAR LETTER DDHA
100F	ထ	MYANMAR LETTER NNA
1010	ဒ	MYANMAR LETTER TA
1011	ဓ	MYANMAR LETTER THA
1012	ဇာ	MYANMAR LETTER DA
1013	ဇာ	MYANMAR LETTER DHA
1014	ဇာ	MYANMAR LETTER NA
1015	ပ	MYANMAR LETTER PA
1016	ဖ	MYANMAR LETTER PHA
1017	ဘ	MYANMAR LETTER BHA
1018	ဗ	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	ဥ	MYANMAR LETTER MON II
1025	ဧ	MYANMAR LETTER U
1026	ဨ	MYANMAR LETTER UU ≡ 1025 ဥ 102E ဣ
1027	ဣ	MYANMAR LETTER MON E
1028	■	<RESERVED>
1029	ဣ	MYANMAR LETTER O
102A	ဣ	MYANMAR LETTER AU
102B	■	<RESERVED>

Dependent vowel signs

102C	ာ	MYANMAR VOWEL SIGN AA
102D	ိ	MYANMAR VOWEL SIGN I
102E	ီ	MYANMAR VOWEL SIGN II
102F	ု	MYANMAR VOWEL SIGN U
1030	ူ	MYANMAR VOWEL SIGN UU

Myanmar(Shan)

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 ၊ dev anagari danda
104B	။	MYANMAR SIGN SECTION → 0965 ။ dev anagari double danda

Various signs

104C	ၵ	MYANMAR SYMBOL LOCATIVE
104D	ၶ	MYANMAR SYMBOL VOMPLETED
104E	ၷ	MYANMAR SYMBOL AMFOREMETION
104F	ၸ	MYANMAR SYMBOL GENITIVE

Pali and Sanskrit extensions

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

Shan extensions

1066	န	SHAN LETTER MAI PAK NGA
1067	န	SHAN LETTER MAI NGA
1068	ဉ	SHAN LETTER MEDIAL VA
1069	ဉ	SHAN LETTER MEDIAL PHA
106A	ေ	SHAN LETTER AOU
106B	ိ	SHAN LETTER A
106C	ီ	SHAN LETTER AY
106D	ံ	SHAN LETTER AI
106E	ံ	SHAN LETTER AIM
106F	ံ	SHAN TONE MARK ONE
1070	ံ	SHAN TONE MARK TWO
1071	ံ	SHAN TONE MARK THREE
1072	ံ	SHAN TONE MARK FOUR
1073	ံ	SHAN TONE MARK FOUR

ANNEX-5
109F

1000

Myanmar(Karen)

	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	ထ 1011	အ 1021	ေ 1031	၁ 1041	ဖ 1051				
2	ဂ 1002	ဒ 1012		ဲ 1002	၂ 1042	ဗ 1052				
3	ဃ 1003	မ 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	ဥ 1025		၅ 1045	ဇ 1055	၈ 1065	၉ 1075		
6	ဆ 1006	ဖ 1016	ဥ 1026	ံ 1036	၆ 1046	ာ 1056		ံ 1076		
7	ဇ 1007	ပ 1017	ဇ 1027	့ 1037	၇ 1047	ာ 1057		ံ 1077		
8	စ 1008	တ 1018		း 1038	၈ 1048	ွ 1058		ံ 1078		
9	ဥ 1009	မ 1019	သ 1029		၉ 1049	ွ 1059		း 1079		
A	ည 100A	ယ 101A	သွ 102A		I 104A			ံ 107A		
B	ဇ 100B	ရ 101B			II 104B			ံ 107B		
C	ဥ 100C	လ 101C	ာ 102C		၂ 104C			ံ 107C		
D	ဥ 100D	ဝ 101D	ံ 102D		၃ 104D					
E	ပ 100E	သ 101E	ံ 102E		၄ 104E					
F	က 100F	ဟ 101F	ူ 102F		၅ 104F					

ANNEX-2

1000

Consonant

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	ဉ	MYANMAR LETTER NNYA
100B	တ	MYANMAR LETTER TTA
100C	ထ	MYANMAR LETTER TTHA
100D	ဒ	MYANMAR LETTER DDA
100E	ဌ	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	ဘ	MYANMAR LETTER BHA
1018	ပ	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	ဇ	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	ဥ	MYANMAR LETTER MON II
1025	ဣ	MYANMAR LETTER U
1026	ဣ	MYANMAR LETTER UU ≡ 1025 ဥ 102E ဣ
1027	ဣ	MYANMAR LETTER MON E
1028		<RESERVED>
1029	ဣ	MYANMAR LETTER O
102A	ဣ	MYANMAR LETTER AU
102B		<RESERVED>

Dependent vowel signs

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 ၊ dev anagari danda
104B	။	MYANMAR SIGN SECTION → 0965 ။ dev anagari 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	၆	MYANMAR LETTER SIGN VOCALIC R
1057	၇	MYANMAR LETTER SIGN VOCALIC RR
1058	၈	MYANMAR LETTER SIGN VOCALIC L
1059	၉	MYANMAR LETTER SIGN VOCALIC LL

Karen extensions

1065	၀	KAREN MEDIAL LA
1074	၁	KAREN LETTER SHAA
1075	၂	KAREN LETTER HAA
1076	၃	KAREN MEDIAL YA
1077	၄	KAREN DEVOWELISER EI THII
1078	၅	KAREN DEVOWELISER A' THII
1079	၆	KAREN DEVOWELISER HPLE HSI
107A	၇	KAREN DEVOWELISER HA THII
107B	၈	KAREN DEVOWELISER KAY PHO
107C	၉	KAREN DEVOWELISER KHLA THII