Title: **Proposal of 4 Myanmar Semivowels**
Source: **Myanmar Unicode and NLP Research Center**
Status: **National Contribution**
Action: **For consideration by WG2**
Date: 21-June-2004

L2/04-273

This document requests additional characters to be added to the UCS within Myanmar block. It also requests to pull 4 semivowel signs out of *virama model.* This contains the proposal summary form.

A. Administrative

1. Title
Proposal of 4 Myanmar semivowels

2. Requester's name
Myanmar Unicode & NLP Research Center

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

4. Submission date
2004-06-21

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
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.

1b. Name of the existing block
Myanmar (U+1000 to U+109F)

2. Number of characters in proposal
4

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

ISO_semivowels5.doc

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
Sample Myanmar documents on http://myanmars.net/unicode/

5a. Is a repertoire including character names provided?
Not yet.

5b. If YES, are the names in accordance with the character naming guidelines in Annex L of ISO/IEC 10646-1: 2000?

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?
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: zhtut@myanmars.net, myanmar-nlp@mail4u.com.mm
internet: 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, see below.

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.)?

ISO_semivowels5.doc

Yes.

2b. If YES, with whom?
- Myanmar IT Standardization Committee (national body)
- Myanmar Computer Federation
- 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://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 example documents attached.

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

5b. If YES, where?
In Myanmar (formerly Burma)

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?
Still yet to be confirmed soon.

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.

ISO_semivowels5.doc

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

The 4 semivowels being submitted to add now were also widely recognized as conjunct consonants, written in a different form. However these 4 semivowels functions more as a semivowel, than a conjunct consonant.

Since it was believed to be the same character, which appears differently, they are not encoded in the UCS before. Now with reliable proofs and papers, this issue is raised again to be properly encoded in UCS.

D.2. User Community

There are over 50 million people residing in Myanmar and living abroad. They do speak, read, and write Myanmar language.

D.3. Proposed Characters

These are the 4 semivowel signs now being submitted:
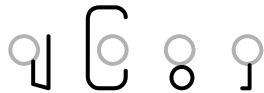
D.4. Character Properties of Proposed Characters

The 4 proposed characters generates up to 11 semivowels, in combination such as:

D.5.1 Reasons why 4 semivowels has to be encoded

1. A medial is a consonant whose form has metamorphosed into a particular symbol and combines with an initial consonant. In this manner one, two or three medials may be attached to a consonant.

2. There are four consonants in the Myanmar alphabet that are used as medials. These are ya, ra, wa and ha. When these assume the role of medials their forms become ျ , ြ , ွ and ှ respectively. These four consonants are remarkable in that, in addition to being consonants, they also exhibit characteristics of vowels and thus they are also termed semi-vowels. There may one, two or three medial symbols combined to a consonant.

3. "The Myanmar orthography" published by the Myanmar Language Commission in October 1986, which the Ministry of Education had prescribed by notification as the national standard mentions eleven medial symbols as described below:-

   ျ   ya pinh ('uplifting ya')

   ြ   ra rit ('encircling ra')

   ွ   wa hswei ('subscript wa')

   ှ   ha htow ('subscript ha')

   ျွ   ya pinh -wa hswei ('uplifting ya with subscript wa')

   ြွ   ra rit -wa hswei ('encircling ra with subscript wa')

   ျှ   ya pinh -ha htow ('uplifting ya with subscript ha)

   ြှ   ra rit –ha htow ('encircling ra with subscript ha')

   ွှ   wa hswei –ha htow ('subscripts wa & ha in combination')

   ျွှ   ya pinh –wa hswei –ha htow ('uplifting ya with subscripts wa & ha')

   ြွှ   ra rit –wa hswei –ha htow ('encirling ra with subscripts wa & ha')

4. It can be seen that there are the four basic medial symbols, five possible combinations involving two basic symbols and two combinations involving three symbols. Although 'ya pinh' is but 'ya pet let' (i.e. 'supine ya') (also known as 'ya-nge' or small ya') in another guise, 'ra rit' is but 'ra –gauk' (i.e. curved ra'), 'wa hswei' is but 'wa' in a subscript position and 'ha htow' is but 'ha' in another guise, the initial and medial consonants cannot be pronounced separately once they have been combined. The following examples show that the difference in pronunciation as well as in meaning of these two forms.

ကျား *(kya:)* is not pronounced *ka ya:* (tiger ≠ kayah) (ကယား)

ကြား *(kya:)* is not pronounced *ka r(y)a:* (hear ≠ beaker/kettle) (ကရား)

ကွေ *(kway)* is not pronounced *ka way* (veer; part ≠ wizard) (ကဝေ)

ချ *(khya)* is not pronounced *kha ya* (hit;put down ≠ be obsequious) (ခယ)

ခြာ *(khya)* is not pronounced *kha r(y)a* (central ≠ sea holly) (ခရာ)

ခွာ *(khwa)* is not pronounced *kha wa* (hoof ≠ laundryman) (ခဝါ)

ခွဲ *(khwe)* is not pronounced *kha we* (divide ≠ luffa plant or its fruit) (ခဝဲ)

ပွာ *(pwa)* is not pronounced *pa wa* (suggest expansion ≠ scarf) (ပဝါ)

ဖြေ *(hpye)* is not pronounced *hpa r(y)e* (widen/stretch out ≠ watermelon) (ဖရဲ)

ဖွား *(hpwa:)* is not pronounced *hpa wa:* (be born ≠ pal/sole) (ဖဝါး)

များ *(mya:)* is not pronounced *ma ya:* (many ≠ wife) (မယား)

မြ *(mya)*　　　is not pronounced *ma r(y)a* (emerald ≠ not get) (မရ)

မြေ *(mye)* is not pronounced *ma r(y)e* (fixed/fast ≠ not brave) (မရဲ)

မြော *(myaw)* is not pronounced *ma r(y)aw* (be in a coma ≠ do not mix) (မရော)

မွေး *(mway:)* is not pronounced *ma way:* (give birth ≠ not far) (မဝေး)

မွဲ *(mwe)* is not pronounced *ma we* (look dull ≠ not have an accent) (မဝဲ)

From the above it can be seen that by separating the medial form the initial and placing it as the second consonant results in a change in pronunciation as well as an altogether different meaning. It may be noted that by combining with a medial results in the elision of the vowel "*a*" inherent in the initial consonant. Results are similar in the case of consonants combining with either two or three medials.

A feature of the Pali script is the employment of conjunct consonants. These are variously known as dvebhava (doubled) or samyoga (united) consonants

(a) In form, romanised Pali melds the consonants in a horizontal direction, while Pali written in the Myanmar script does this by means of a subscript. However both methods share the following similarities :-

(1). Conjunct consonants always come after a vowel or a consonant with an inherent vowel; (E.g. atta = အတ္တ; khetta = ခေတ္တ)

(2). The consonant following the initial consonant is devowellized. This is evident in the above romanised example where the first 't' has its inherent vowel elided, while the word written in Myanmar script is also pronounced အတ်တ၊ ခေတ်တ although no devowelizer or "killer" sign appears on the first or upper တ in both conjuncts.

(3). This is common in principle with the Virama in Devanagari.

(b) The Myanmar language has borrowed and assimilated many words from Pali due to its long association with it. Some are used in original form as it is in the case of the Pali words given as examples in the above paragraph 6.(a).(1) but many Pali words employing conjunct consonants have been adopted with a devowellized final. (e.g. Pali santāna > သန္တာန် in Myanmar; Pali nikāya > နိကာယ် in Myanmar etc.) There are also other methods of adapting Pali words.

(c) This style of assimilating Pali conjunct consonants has been extended to writing some Myanmar words in a similar fashion as show below:-

ယောက္ခမ၊ ရမ္မက်၊ လိမ္မာ၊ သေတ္တာ

(d) Moreover, such conjunct consonants can also be found in words adapted from English, and from other languages as well. For example

ကုမ္ပဏီ < company;	အစ္စရေး < Israel;
ပါကစ္စတန် < Pakistan;	ထရန်စစ္စတာ < transistor;

(e) It has been described in above paragraph 6 (a) (1) and illustrated in the various examples given in the following sub- paragraphs that conjunct consonants always come after an initial vowel or a consonant with an inherent vowel. This is not the case with the medials where a medial symbol or symbols may combine with a single consonant or with consonants any where within a word. A few examples of words that are complete in themselves constituting just one consonant combined with a medial are given in paragraph 5. Many more words of a similar nature or more complex compounds and even combinations constituting an initial & medial followed by conjunct-consonants (such as မြန္မာ၊ လှဒန္ဒ္ရိန်) or a series of medial combinations in a word (such as ဖြဗျုဂ္ဏား) are to be found, Thus, it can be seen that the virama model by itself alone cannot substitute the medials although it shares with the medials, the characteristic of eliding the inherent vowel *"a"* in the preceding consonant. A sizeable part of the Myanmar lexicon and proper names would be left out should the medials be not given their proper place in the Myanmar character set on the premise that the virama is capable of performing their duty as well.

(f) Furthermore, since the medials constitute four consonants that are semi-vowels these display characteristics such as altered forms as well as a change in pronunciation; these should be treated as a separate and distinct development.

(g) It is therefore proposed that the four medials be included in the character set together with the capability to form various possible combinations within themselves in addition to the virama model needed for conjunct consonants in the lexicon that may or may not be Pali derivatives.

5. The Medials in the Myanmar language seem to be a unique development and each of them makes a specific contribution meaningwise when they combine with a consonant either individually or in combination with others.

(a) Shown below are how some of the medials exercise their influence when combined with consonant ka (က)

က ကျ ကြ ကွ

က by itself carries the connotations of 'touching, catching, connecting, joining, combining and binding etc.'

In ကျ it is extended further to mean continually or continuously touching, catching, connecting, joining etc'.

In ကြ it denotes the notion of 'touching, catching, connecting, joining, combining and binding etc with great force and in unison'.

While ကွ carries it to 'the greatest magnitude, denoting, touching, catching, connecting, joining, combining, binding etc. with the greatest force excessively'

(b) Another example will be given below, this time with the word pā: as the root in combinations with some of the medials

ပါး - by itself denotes the notion of 'being carried away, reaching some place or spreading either in a specific direction or to all directions of the compass'
- From the idea of spreading, it follows that sth has been this way and that way so that it will become relatively thinner in comparison to its former thickness.

- The expression လက်ဆောင်ပါးသည် has the literal meaning of 'having a gift to be carried away' (to the person intended) thus reflecting one aspect of its meaning which is causing sth to be carried away.

- The other aspect of denoting thinness is evident in the words ပါးလှစ် which means 'sheer, diaphanous', နည်းပါး meaning 'few' which can be the result of spreading sth too thin and လျော့ပါး which means to diminish, decrease or become less.

- In ပြား, the essential meaning of ပါး i.e. to carry away, spreading in all directions 'is further intensified in the various combinations with ပြား such as ကွဲပြား which means 'distinct' (evidently carrying away sth from another results in separation, and as such may be said to become distinct) ေထွပြား on the other hand means literally 'diverse and varied', and therefore hints at being complex, or being complicated and များပြား indicates both numerousness and variety. The other meaning of ပြား being flat is naturally related to being thin as denoted by the word ပါး

- In the case of ပွား it has the notion of spreading, radiating, proliferating with မုေလးပွား being a religious offering made of paper ribbons fluffed up like a flower; တိုးပွား meaning to proliferate, thrive'; ပြန့်ပွား 'to spread, disseminate'; ပွားစီး meaning 'to increase, flourish, prosper' and ပွားများ having the meanings of increasing, flourishing (in the material sense) and to nurture some mental aspect (in the spiritual sense)

An interesting development can be discerned in the case of the word ပျား the word for a bee or bees. Although being aware of the existence of 'solitary bees' in general bees are regarded as the social kind which words like 'spelling-bee' and the expression 'a bee-hive of activity' amply illustrate. However, it seems to have been derived from ပျ which is associated with ပျာပျာသလဲ၊ ပျာယာခတ်၊ ပျယီးပျယာ etc. all of which suggests a flurry of motion, or frenzied activity. Nevertheless ပျ itself may be related to ပါး and ပါ in which the inherent meaning of carrying and being carried away are to be found.

D.5.2. Serious conflicts in current encoding algorithm

When the semivowels or medials are encoded as same as stacked characters, the following serious encoding conflicts occur.

According to Unicode 4.0 Standard (page 272):

> For example, the word *krwe* ့ [kjwei] ("to drop off ") would be written via the following sequence:

U+1000 က *ka* + U+1039 ◌် *virama* + U+101B ရ *ra* + U+1039 ◌် *virama* + U+101D ဝ *wa*

+ U+1031 ေ *vowel sign e* → ကြွေ *krwe*

According to Unicode 4.0 Standard (page 273) says:

> *subscript consonant* <U+1039, [U+1000–U+1019, U+101C, U+101E, U+1020, U+1021]> ◌္သ

Therefore, according to Unicode 3 and 4, the following syllables have to be encoded this way:

တျ	=	က	+	virama	+	ယ
		1010		1039		101A
ကြ	=	က	+	virama	+	ရ
		1000		1039		101B
ကွ	=	က	+	virama	+	ဝ
		1000		1039		101D
ကှ	=	က	+	virama	+	သ
		1000		1039		101F

And all stacked consonants this way.
Note that the ***medial characters are being encoded as same as stacked characters***.

က္က = က + virama + က
          1000   1039   1000

က္ခ = က + virama + ခ
          1000   1039   1001

e.g.   စက္ကူ   meaning "paper"

ဒုက္ခ   meaning "trouble"

So far, everything seems to be ok.

BUT, when we encode these old use Myanmar syllables below, there comes serious encoding conflicts here. It has 2 different style of writings. It may or may not be the same meaning. But should it be the same encoding?

You will see the same encoding for different syllables in the examples below:

Pitfalls in virama (U1039) + HA (U101F) combinations

ကဿ = က + ဏ + virama + ဟ
                  1010   100F   1039   101F

ကဿ = က + ဏ + virama + ဟ (old use)
                  1010   100F   1039   101F

Pitfalls in virama + YA (U101A) combinations

ဒိက္ရ = ဒ + ိ + က + virama + ယ
               1012   102D   1010   1039   101A

ဒိက္ရ = ဒ + ိ + က + virama + ယ (old use)
               1012   102D   1010   1039   101A

Pitfalls in virama + RA (U101B) combinations

ပြု = ပ + virama + ရ + ◌ု
          1015   1039   101B   102F

သြု = သ + virama + ရ + ◌ု (old use)
          101E   1039   101B   102F

Pitfalls in virama + SA (U101E) combinations
Apart from those medial characters, there also are a few characters, which has to be separated apart from the virama model.

ပသ္သ = ပ + သ + virama + သ
              1015   101E   1039   101E

လသ္သ = လ + သ + virama + သ
              101C   101E   1039   101E

D.5.3. Storage and performance inefficiency

Virama Model has been applied to be 'economy in coding'. If the 4 semivowel signs can be encoded in the UCS, there would be a far more economy coding method.

As an example, let us encode the name of the country "The Union of Myanmar" as follows:

ပြည်ထောင်စုမြန်မာနိုင်ငံတော်

current ASCII models

ပြည်ထောင်စုမြန်မာနိုင်ငံတော်

bytes used: <u>28 bytes</u> in TXT, , bytes in HTML

<u>according to Unicode 3 & 4</u>
70 bytes have to be used in TXT

ပ + V + ရ + ည + ◌ံ + ZWNJ + ထ + ေ + ာ + င + ◌ံ + ZWNJ + စ + ◌ု + မ + V + ရ + န + ◌ံ + ZWNJ +

မ + ာ + န + ◌ံ + ◌ု + င + ◌ံ + ZWNJ + င + ◌ံ + တ + ေ + ာ + ◌ံ + ZWNJ

224 bytes have to be used in HTML

ပ + <mark style="background:red">္</mark> + ရ + ည + ္ + <mark>‌</mark> + ထ + ေ + ာ +
င + ္ + <mark>‌</mark> + စ + ု + မ + <mark style="background:red">္</mark> + ရ + န +
္ + <mark>‌</mark> + မ + ာ + န + ိ + ု + င + ံ +
တ + ေ + ာ + ္ + <mark>‌</mark>

<u>proposing model</u>
only 56 bytes have to be used in TXT

ပ + ြ + ည + ◌ံ + ZWNJ + ထ + ေ + ာ + ◌ံ + င + စ + ◌ု + မ + ြ + န + ◌ံ + ZWNJ + မ + ာ + န + ◌ံ + ◌ု + ◌ံ + Z

WNJ + င + ◌ံ + တ + ေ + ာ + ◌ံ + ZWNJ

only 210 bytes have to be used in HTML

ပ + <mark style="background:lime">&#4154</mark>; + ည + ္ + <mark>‌</mark> + ထ + ေ + ာ + င +
္ + <mark>‌</mark> + စ + ု + မ + <mark style="background:lime">်</mark> + န + ္ + <mark>‌</mark> +
မ + ာ + န + ိ + ု + င + ံ + တ + ေ +
ာ + ္ + <mark>‌</mark>

Thus, the proposing model (with the semivowel signs encoded) will be saving in rough average of 20% in normal text, and about 5% in HTML text. A lot of harddisk space for all Myanmar users.

E. Bibliography

1. Myanmar-Myanmar-English Dictionary
   by Myanmar Language Commission
2. Myanmar Orthography
   by Myanmar Language Commission, 1999 Edition
3. Myanmar-English Dictionary
   by Myanmar Language Commission
4. Voharatthapakathani|
   by Rev. Kyaw Aung San Htar
5. Brahma Niruttinayasangaha
   by Rev. U Kosalla
6. Voharalinattha Dipani
   by U Chein
7. Myanmar Saddanaya and Atthanaya
   by U Po Sein
8. The Unicode Standard 3.0
9. The Unicode Standard 4.0
   by Unicode Consortium, 2002
10. Collection of Myanmar NLP Research Work and Papers
    http://myanmars.net/unicode/

ANNEX-1

1000 — Myanmar — 109F

	100	101	102	103	104	105	106	107	108	109
0	1000	1010	1020	1030	1040	1050				
1	1001	1011	1021	1031	1041	1051				
2	1002	1012		1002	1042	1052				
3	1003	1013	1023		1043	1053				
4	1004	1014	1024		1044	1054				
5	1005	1015	1025		1045	1055				
6	1006	1016	1026	1036	1046	1056				
7	1007	1017	1027	1037	1047	1057				
8	1008	1018		1038	1048	1058				
9	1009	1019	1029	1039	1049	1059				
A	100A	101A	102A	103A	104A					
B	100B	101B		103B	104B					
C	100C	101C	102C	103C	104C					
D	100D	101D	102D	103D	104D					
E	100E	101E	102E		104E					
F	100F	101F	102F		104F					

1000   **Myanmar**   **1065**

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	ဉ	MYANMAR LETTER NYA
100A	ည	MYANMAR LETTER NNYA
100B	ဋ	MYANMAR LETTER TTA
100C	ဌ	MYANMAR LETTER TTHA
100D	ဍ	MYANMAR LETTER DDA
100E	ဎ	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 ◌ီ

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
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)

SemiVowel

103A	◌ျ	MYANMAR SEMIVOWEL YA
103B	◌ြ	MYANMAR SEMIVOWEL RA
103C	◌ွ	MYANMAR SEMIVOWEL WA
103D	◌ှ	MYANMAR SEMIVOWEL HA

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	၎င်း	MYANMAR SYMBOL AMFOREMETION
104F	၏	MYANMAR SYMBOL GENITIVE

Pali and Sanskrit extensions

1050	ၐ	MYANMAR LETTER SHA
1051	ၑ	MYANMAR LETTER SSA
1052	ၒ	MYANMAR LETTER VOCALIC R
1053	ၓ	MYANMAR LETTER VOCALIC RR
1054	ၔ	MYANMAR LETTER VOCALIC L
1055	ၕ	MYANMAR LETTER VOCALIC
1056	ၖ	MYANMAR VOWEL SIGN
1057	ၗ	MYANMAR VOWEL SIGN
1058	ၘ	MYANMAR VOWEL SIGN
1059	ၙ	MYANMAR VOWEL SIGN VOCALIC LL