

Medial Disunification

I am happy to hear that efforts to accommodate minority languages spoken in Burma are being addressed by the committee. It is the sole reason that my earlier support for Karen, Mon and Shan have not been brought forward yet in a commercial product.

While I welcome this, I am not sure that the current recommendation is the best method — the transcoding required is an unnecessary burden as the medials can already be handled properly.

New nominal glyphs will be needed to address Mon, Shan, Karen and other minority languages. Rather than introduce an exceptional glyph (the MYANMAR SIGN VIRAMA à la the Khmer encoding), it seems that a simplification of the medials can be achieved in one of two other ways that will avoid transcoding issues:

1. Add new nominal glyphs to represent all base consonants for minority languages that require different shaping rules.

This will need to be done to handle other characters, e.g., Shan /*pha*/, Mon /*nga*/). It strikes me that there is no reason why there can't be a S'gaw Karen ဝ် /*ya*/, a Mon ဝ် /*la*/, for example, that then have their own unique rendering, even if their nominal glyph shape is identical or similar to a Burmese nominal glyph. This avoids the need to transcode existing Burmese and Pali texts that have been produced in the past few years.

2. Add new nominal glyphs to represent only those base consonants for minority languages that have unique glyph shapes and use shaping rules to display language-specific forms whenever necessary.

In this scenario, items like S'gaw Karen ဝ် /*ya*/ and Mon ဝ် /*la*/ would use the existing Burmese glyphs ဝ် and ဝ် but incorporate language-specific tables to produce the necessary subscript forms (ကျ ဝ် ဝ် ဝ်). This too avoids the transcoding problems.

In either instance the necessary forms can be produced. E.g:

| | | |
|----|------------------|---|
| ကျ | KA + VIRAMA + LA | currently implemented as a language feature |
| ဝ် | KA + VIRAMA + YA | implemented as a historical ligature |
| ဝ် | KA + VIRAMA + SA | implemented as a historical ligature |
| ဝ် | KA + VIRAMA + HA | implemented as a historical ligature |

Disunification of TALL AA (က/ဝါ)

I find the rationale for disunifying these two glyphs into separate Unicode entities (as explained in the current proposal) unnecessary. A better option might be to encode a separate Karen -aa vowel glyph that undergoes no transformations and to maintain the existing Burmese -aa which would continue to be transformed as appropriate. This option avoids transcoding existing Burmese texts and allows the two vowel glyphs to remain distinct.

Disunification of GREAT SA (သ)

I am not aware of any set of circumstances that require this disunification. This glyph is used in Pali and Pali-derived words traditionally. I am in agreement that the syllable /*ssa*/ is by convention not written as သ — except in Pyu inscriptions and the desire to represent these forms in a modern Burmese orthography. An attempt to disunify them is not supported by any compelling reason as the need to represent the three in isolation can be accomplished through

other means, i.e., SA + VIRAMA + SA [သ]; SA + VIRAMA + ZWNJ + SA [သ်သ]; SA + VIRAMA + (ZWJ) + SA [သ့] though I have implemented this last one as a historical ligature.

All 3 forms presented in n3xxx-myanmar.pdf are in exclusive distribution. Creating a model that makes the default representation of SA+VIRAMA+SA a historical form [သ့] that is only used in exceptional circumstances seems awkward and wholly unnecessary.

Disunification of ASAT

If a comparison is to be made between ASAT, a new Burmese COENG and the existing Khmer COENG, it should be remembered that this artificial mark was highly controversial and continues to be in some circles. Its inclusion should be warranted by more than just simplification.

Representation of KINZI

KINZI is indeed difficult to render correctly but not impossible. If, as it appears, the only reason for this modified representation is as the result of the overall simplification of the encoding model, it should be recognized that the modified representation is *not* required for the accurate display of KINZI.

Although a simplification of KINZI would be welcome, the reference to a distinction between modern သခါရ and the older သဋ္ဌာရ as well as the unattested form သဋ္ဌိာရ is moot under the current model, for reasons mentioned earlier. E.g.:

| | |
|---------|--|
| သခါရ | KA + NGA + VIRAMA + KHA + AA + RA |
| သဋ္ဌာရ | KA + NGA + ZWNJ + VIRAMA + KHA + AA + RA |
| သဋ္ဌိာရ | KA + NGA + ZWNJ + VIRAMA + ZWNJ + VIRAMA + KHA + AA + RA |
| ငြင်း | NGA + VIRAMA + RA + NGA + VIRAMA + VISARGA |

Change of glyph for SYMBOL AFOREMENTIONED

I concur entirely with this change.

Figure 5

While I cannot say where the information for this chart originated, I do note that it mirrors my pre-Unicode implementation of Burmese from the early- to mid-90s. Lloyd Anderson (of Ecological Linguistics) and I had dealings concerning many minority scripts and he offered several of my implementations in the early 90s as retail products for Macintosh users.

Impact on current implementations

It has been stated that there is only one complete Unicode 4.1 compliant implementation which is not entirely accurate. I have had an implementation of Burmese available since March of 2003 (the development of which began in November of 2002) based on Unicode 3.0 and while I have not had an opportunity to verify that I have addressed all changes that may be required in 4.1, several enhancements were added in October of 2004 and again in April of 2005 towards that end.

While the number of official users is relatively small, there is reason to believe that my fonts have a larger user-base. They represent Burmese language scholars as well as native speakers and learners and at least one Pali scholar (and I am currently working on a Pyu implementation using the existing Burmese encoding). These proposed changes will necessitate a large number of transcoding problems. While some of these may be inevitable, it is hard to justify additional development time simply because of limitations in other rendering mechanisms.

For a short time the Macintosh was the more prevalent operating system used in Burma by publishers of a variety of minority languages. That has changed to cheaper, alternative systems but it is somewhat alarming to me that Unicode is now being subjected to changes required by inadequate rendering engines, particularly since there is at least one reasonable alternative methodology. As a developer supporting minority scripts for more than 15 years, I have to hope with some trepidation that this is not the beginning of an unfortunate turn of events where insufficient rendering engines dictate the future of Unicode encodings.

Sample Burmese Text

အလောင်းမင်းတရား

မဟာဓမ္မရာဇာဓိပတိလက်ထက်တွင် အင်းဝနေပြည်တော်ကြီးသည် မငြိမ်မသက်ဖြစ်နေသည်။ မဏိပူရ ကသည်းများက အင်းဝနေပြည်တော်၏ မြောက်ဘက်တလွှားကို တိုက်ခိုက်ဖျက်ဆီးသည်။ အောက်မြန်မာနိုင်ငံ ဟံသာဝတီသားတို့ကလည်း ပုန်ကန်သည်။ မတ္တရာအုတ်ဖိုရို ကွေ့ရှမ်းများကလည်း ထကြွသည်။

အင်းဝကို သိမ်းပိုက်ပြီးသော ဟံသာဝတီတပ်များသည် မြောက်ဘက်တလွှားကျေးရွာများကို သစ္စာခံခိုင်းရန် လာကြရာ မုဆိုးဖိုရွာသို့ ရောက်ရှိလာသည်။ တချိန်တည်းမှာပင် ကွေ့ရှမ်းများကလည်း သစ္စာခံခိုင်းရန် ရောက်ရှိလာသည်။ ဦးအောင်ဇေယျသည် အဖွဲ့နှစ်ဖွဲ့ကို ခြေငံစွာ ဆက်ဆံသည်။ မည်သူ့သစ္စာကိုမျှ ခံယူခြင်းမပြုပေ။ ဟံသာဝတီတပ်ဖွဲ့ကို အပြန်ခရီးတွင် လမ်းမှဖြတ်၍ တိုက်ခိုက်သည်။ ဟံသာဝတီတပ်ဖွဲ့များ အထိအခိုက်အကျပုအဆုံးများစွာဖြင့် ပြန်ရသည်။

မဟာဓမ္မရာဇာဓိပတိလက်ထက်တွင် အင်းဝနေပြည်တော်ကြီးသည် မငြိမ်မသက်ဖြစ်နေသည်။ မဏိပူရ ကသည်းများက အင်းဝနေပြည်တော်၏ မြောက်ဘက်တလွှားကို တိုက်ခိုက်ဖျက်ဆီးသည်။ အောက်မြန်မာနိုင်ငံ ဟံသာဝတီသားတို့ကလည်း ပုန်ကန်သည်။ မတ္တရာအုတ်ဖိုရို ကွေ့ရှမ်းများကလည်း ထကြွသည်။

အင်းဝကို သိမ်းပိုက်ပြီးသော ဟံသာဝတီတပ်များသည် မြောက်ဘက်တလွှားကျေးရွာများကို သစ္စာခံခိုင်းရန် လာကြရာ မုဆိုးဖိုရွာသို့ ရောက်ရှိလာသည်။ တချိန်တည်းမှာပင် ကွေ့ရှမ်းများကလည်း သစ္စာခံခိုင်းရန် ရောက်ရှိလာသည်။ ဦးအောင်ဇေယျသည် အဖွဲ့နှစ်ဖွဲ့ကို ခြေငံစွာ ဆက်ဆံသည်။ မည်သူ့သစ္စာကိုမျှ ခံယူခြင်းမပြုပေ။ ဟံသာဝတီတပ်ဖွဲ့ကို အပြန်ခရီးတွင် လမ်းမှဖြတ်၍ တိုက်ခိုက်သည်။ ဟံသာဝတီတပ်ဖွဲ့များ အထိအခိုက်အကျပုအဆုံးများစွာဖြင့် ပြန်ရသည်။

မဟာဓမ္မရာဇာဓိပတိလက်ထက်တွင် အင်းဝနေပြည်တော်ကြီးသည် မငြိမ်မသက်ဖြစ်နေသည်။ မဏိပူရ ကသည်းများက အင်းဝနေပြည်တော်၏ မြောက်ဘက်တလွှားကို တိုက်ခိုက်ဖျက်ဆီးသည်။ အောက်မြန်မာနိုင်ငံ ဟံသာဝတီသားတို့ကလည်း ပုန်ကန်သည်။ မတ္တရာအုတ်ဖိုရို ကွေ့ရှမ်းများကလည်း ထကြွသည်။

အင်းဝကို သိမ်းပိုက်ပြီးသော ဟံသာဝတီတပ်များသည် မြောက်ဘက်တလွှားကျေးရွာများကို သစ္စာခံခိုင်းရန် လာကြရာ မုဆိုးဖိုရွာသို့ ရောက်ရှိလာသည်။ တချိန်တည်းမှာပင် ကွေ့ရှမ်းများကလည်း သစ္စာခံခိုင်းရန် ရောက်ရှိလာသည်။ ဦးအောင်ဇေယျသည် အဖွဲ့နှစ်ဖွဲ့ကို ခြေငံစွာ ဆက်ဆံသည်။ မည်သူ့သစ္စာကိုမျှ ခံယူခြင်းမပြုပေ။ ဟံသာဝတီတပ်ဖွဲ့ကို အပြန်ခရီးတွင် လမ်းမှဖြတ်၍ တိုက်ခိုက်သည်။ ဟံသာဝတီတပ်ဖွဲ့များ အထိအခိုက်အကျပုအဆုံးများစွာဖြင့် ပြန်ရသည်။