

ISO/IEC JTC1/SC2/WG2 N3069

2006-04-06

Doc Type: Working Group Document

Title: Concerns Regarding WG2 N3043R, Myanmar Additions to 10646

Source: Unicode Consortium

Status: Liaison Statement

Action: For consideration by JTC1/SC2/WG2

Date: 2006-04-06

The Unicode Consortium has become aware of a document submitted to WG2 (WG2 N3043R) which proposes addition of seven Myanmar characters to the UCS.

Specifically, this proposal is being presented as an urgent request, for "fast-tracking" into 10646, and it is our understanding that it is likely that at least one national body will be requesting that these additions be added to FPDAM 3, in the context of ballot comment resolution of PDAM 3.

The document in question was submitted to WG2 in March, after the last opportunity for the UTC or the U.S. national body to review such submissions formally had passed. (The regularly scheduled joint meeting was held February 6 - 9, 2006.)

Ordinarily, in the case of documents submitted shortly before a WG2 meeting requesting the addition of a few characters to an existing script or symbol set, national body delegates may review such proposals on their merit and consider them appropriate for late inclusion in a ballot, knowing that the national bodies (and the UTC) will have another ballot period to review such additions.

However, in the case of the proposed seven Myanmar additions, the situation is not ordinary, but rather extraordinary.

Expert review during the last few weeks has turned up a number of significant points:

- * The seven character additions are not mere character additions, but character disunifications, as the proposal itself notes. Character disunifications are subject to additional constraints by the WG2 Principles and Procedures, and should, in any case, receive extra careful review.
- * Furthermore, while implicit in WG2 N3043R, it may not be clear to all national body reviewers on first review, that Myanmar text encoded by the current UCS would no longer be correctly coded by the UCS once these characters are added, and any existing data would essentially be invalidated by the proposed change.
- * It has become clear that the proposers of the additional Myanmar characters did not consult sufficiently (or at all) with major software vendors before submitting WG2 N3043R, and since

the posting of WG2 N3043R, a number of the commercial vendor experts have voiced serious concerns about both the details of the proposed additions and the implications of the changes for the stability of the standard. Because of the major impact of this proposal, it is vital that major software providers are consulted and given sufficient time to review the proposal.

- * Contrary to the claims made in WG2 N3043R, more than one implementation of the current encoding for Myanmar does exist, either belying or calling into question some of the claims made about the inadequacy of the current set of encoded characters for Myanmar.

Under these circumstances, the Unicode Consortium is of the opinion that it is both technically and procedurally risky to attempt to rush the acceptance of this particular proposal, and in particular to push it immediately into an existing FPDAM document without a due process of deliberation by all the national bodies and commercial vendors who may be interested or impacted.

Judging from the expert feedback we have heard so far, it is clear that there is not yet consensus about the proposal outside of the authors cited in the proposal itself.

Characters are added to the standard in order to allow text to be legibly represented that could not be represented without the change. In the case of complex scripts, where the glyphs are subject to contextual shaping, establishment of the precise requirements often need considerable research and analysis. The Unicode Consortium and national bodies must be confident that the requirements cannot be met by other mechanisms, such as language-sensitive font technology or the use of variation selectors.

They must also be confident that the requirements cannot be met simply by adding new characters in a way that does not establish an incompatible change. Any prospect of a change which would invalidate existing data and/or implementations requires extremely careful review by members of the Unicode Consortium and interested national bodies before there would be confidence that such changes were in fact warranted.