

Title: US / Unicode Consortium Comments on N3611

Date: 2009-4-21

Source: US national body & Unicode Consortium

Status: Joint NB, Liaison contribution

Action: For review by WG2 experts

Distribution: WG2

In N3611, Korea requests the addition of text in Clause 26.1 of ISO/IEC 10646 that will make reference to the Korean national standard KS X 1026-1:2007. Their proposal is to add the following text:

Hangul text can be represented in several different ways in this standard. For information exchange, to ensure interoperability, Korean Standard KS X 1026-1 (Information Technology – Universal Multiple - Octet Coded Character Set(UCS) - Hangul - Part1 Hangul processing guide for information interchange) must be followed.

The US national body and the Unicode Consortium understand Korea's desire to promote the use of particular representations for Hangul text and desire to work together with Korea toward that objective. There are, however, some outstanding concerns regarding certain parts of KS X 1026-1:2007 as it relates to normalization defined in Unicode and in ISO/IEC 10646. For that reason, we feel it would be inappropriate to add this particular text to Clause 26.1 at this time. We wish to work together with Korea to resolve the outstanding issues regarding normalization and, in turn, to find the best ways to promote appropriate recommendations for interchange of Hangul text.

Background

During WG2 meeting 52 (Redmond), Korea gave an introduction to the recently-published Korean standard KS X 1026-1:2007, and invited NBs to provide feedback. Action item AI-52-11.d was recorded for national bodies and liaison organizations to review the Korean standard and provide feedback to Korea.

In N3536, the US national body and the Unicode Consortium provided feedback expressing "serious concerns about incompatibilities with the important SC2 standard ISO/IEC 10646 and the synchronized Unicode Standard" specifically in relation to normalization. We have received no response from Korea, and those concerns remain very real.

Section 6 of KS X 1026-1:2007 discusses normalization in relation to Hangul text. It makes explicit reference to normalization as defined in UAX#15, which is a normative part of Unicode and of ISO/IEC 10646. It then gives recommendations to use alternate algorithms for normalization of Korean text.

These recommendations, at least as expressed in the English version, are a serious concern because of their incompatibility with the Unicode and ISO/IEC 10646 standards. Without having ability to review the original Korean text of KS X 1026-1 or to review the conformance tests described in KS X 1026-2, it is not clear if the actual intent might have been misunderstood. For instance, it may be the case that Korea's intent was not to change results when normative application of normalization must be applied, but rather generally to recommend the interchange format for Korean text in scenarios in which application of normalization is not normatively required. If that were the case, then a recommendation regarding alternate representations in contexts in which normalization is not required might be possible

and might not lead to these concerns. However, the English text of Section 6 in KS X 1026-1, as currently written, does convey a recommendation to change normalization, and so is problematic.

As stated above, the US national body and the Unicode Consortium are open to working with Korea to find appropriate ways to promote the use of particular encoded representations for use in interchange of Hangul text. Until the concerns regarding normalization in KS X 1026-1, Section 6 are resolved, however, reference to KS X 1026-1:2007 within either Unicode or ISO/IEC 10646 must be seen as problematic.

Therefore, the US national body and the Unicode Consortium would encourage Korea to amend Section 6 of KS X 1026-1. Then, if the existing concerns with that section are resolved, it will be possible to reconsider possible additions in Clause 26.1 that reference KS X 1026-1.