

Universal Multiple Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation
Международная организация по стандартизации

Doc Type: Working Group Document

Title: A response to AI-52-8

Source: Kim, Kyongsok, Head of Delegation

Status: National Position of Rep. of Korea

Action: For consideration by JTC1/SC2/WG2

Date: 2008-10-14

1. This document is a response to an action item AI-52-8 decided at WG2 meeting #52 at Redmond, WA. USA.

AI-52-8 Korea (RoK) (Prof. Kyongsok Kim)

To take note of and act upon the following items.

a. To prepare a contribution elaborating on the differences between Unicode and ISO/IEC 10646 reported in document N3422, and to gather and report on feedback on this document from national bodies and liaison organizations

2. We have not received any feedback RE: this action item.

3. The difference between ISO/IEC 10646:2003 (+ Amd3 and Amd4) and Unicode 5.0

Consider representing a syllable-final letter KIYEOK (U11A8) alone.

1) According to ISO/IEC 10646, two code positions will be needed:
"U115F, U11A8".

2) However, according to Unicode, three code positions will be needed:
"U115F, U1160, U11A8". Note that U1160 is inserted here.

- Korea's proposal as suggested at WG2 meeting in Seoul in 1992 was exactly reflected in ISO/IEC 10646-1:1993.

- The relevant portions in ISO/IEC 10646 and Unicode are shown below in 3.1 and 3.2, respectively.

3,1 The relevant portion in ISO/IEC 10646:2003 + Amd3 + Amd4:

- a syllable-final letter alone =

code positions of syllable-initial filler (U115F) + syllable final letter

- Each of Amd3 and Amd4 changes clause 26.1 slightly. However, those changes do not matter in this regard.

26.1 Hangul syllable composition method

20

In rendering, a sequence of Hangul Jamo (from HANGUL JAMO block: 1100 to 11FF) is displayed as a series of syllable blocks. Jamo can be classified into three classes: Choseong (syllable-initial ~~character~~), Jungseong (syllable-peak ~~character~~), and Jongseong (syllable-final ~~character~~). A complete syllable block is composed of a Choseong and a Jungseong, and optionally a Jongseong.

AMD3

An incomplete syllable is a string of one or more characters which does not constitute a complete syllable (for example, a Choseong alone, a Jungseong alone, a Jongseong alone, or a Jungseong followed by a Jongseong). An incomplete syllable which starts with a Jungseong or a Jongseong shall be preceded by a CHOSEONG FILLER (0000 115F). An incomplete syllable composed of a Choseong alone shall be followed by a JUNGSEONG FILLER (0000 1160).

~~The implementation level 3 shall be used for the Hangul syllable composition method.~~

AMD4

NOTE 1 – Hangul Jamo are not combining characters.

NOTE 2 – When a combining character such as HANGUL SINGLE DOT TONE MARK (0000 302E) is intended to apply to a sequence of Hangul Jamo it should be placed at the end of the sequence, after the Hangul Jamo character which completes the syllable block.

- 3.2 The relevant portion in Unicode 5.0:
- a syllable-final letter alone =
code position of syllable-initial filler (U115F) + syllable-peak filler (U1160) + syllable final letter

120

Transforming into Standard Korean Syllables. A sequence of jamos that do not all match the regular expression for a standard Korean syllable block can be transformed into a sequence of standard Korean syllable blocks by the correct insertion of choseong fillers and jungseong fillers. This transformation of a string of text into standard Korean syllables is performed by determining the syllable breaks as explained in the earlier subsection “Hangul Syllable Boundaries,” then inserting one or two fillers as necessary to transform each syllable into a standard Korean syllable. Thus

$$L [^V] \rightarrow L V_f [^V]$$

$$[^L] V \rightarrow [^L] L_f V$$

$$[^V] T \rightarrow [^V] L_f V_f T$$

where [^X] indicates a character that is not X, or the absence of a character.

121

Examples. In Table 3-13, the first row shows syllable breaks in a standard sequence, the second row shows syllable breaks in a nonstandard sequence, and the third row shows how the sequence in the second row could be transformed into standard form by inserting fillers into each syllable. Syllable breaks are shown by *middle dots* “.”.

Table 3-13. Korean Syllable Break Examples

| No. | Sequence | | Sequence with Syllable Breaks Marked |
|-----|---|---|--|
| 1 | LVTLVVLV _f L _f V _f T | → | LVT · LV · LV · LV _f · L _f V · L _f V _f T |
| 2 | LLTTVVTTVVLLVV | → | LL · TT · VVTT · VV · LLVV |
| 3 | LLTTVVTTVVLLVV | → | LLV _f · L _f V _f TT · L _f VVTT · L _f VV · LLVV |

* * *