Proposal to add nine multi-letter Jamo characters

Kent Karlsson
2008-02-01

Hangul Compatibility Jamo and new Hangul Jamo

TUS 5.0, page 436 says:

---

**Hangul Compatibility Jamo:** U+3130–U+318F

This block consists of spacing, nonconjoining Hangul consonant and vowel (jamo) elements. These characters are provided solely for compatibility with the KS X 1001:1998 standard. Unlike the characters found in the Hangul Jamo block (U+1100..U+11FF), the jamo characters in this block have no conjoining semantics.

The characters of this block are considered to be fullwidth forms in contrast with the halfwidth Hangul compatibility jamo found at U+FFA0..U+FFDF.

**Standards.** The Unicode Standard follows KS X 1001:1998 for Hangul Jamo elements.

**Normalization.** When Hangul compatibility jamo are transformed with a compatibility normalization form, NFKD or NFKC, the characters are converted to the corresponding conjoining jamo characters. Where the characters are intended to remain in separate syllables after such transformation, they may require separation from adjacent characters. This separation can be achieved by inserting any non-Korean character.

- U+200B zero width space is recommended where the characters are to allow a line break.
- U+2060 word joiner can be used where the characters are not to break across lines.

---

I don't think this is the best way to view the Hangul Compatibility Jamo. I think it would be better to view them as Hangul Syllables with Filler.

E.g., U+3131, HANGUL LETTER KIYEOK, and <U+1100, U+1160> (<HANGUL CHOSEONG KIYEOK, HANGUL JUNGSEONG FILLER>) should look and behave (except, unfortunately, for normalisation) just about the same.

Similarly for U+314F, HANGUL LETTER A, and <U+115F, U+1161> (<HANGUL CHOSEONG FILLER, HANGUL JUNGSEONG A>) should look and behave (except, unfortunately, for normalisation) just about the same.

The full list of auxiliary decomposition mappings for compatibility Jamos is as follows in the list below. Note that some of the auxiliary decomposition mappings here refer to proposed characters (L2/07-103), and some even to characters otherwise not yet proposed (but are proposed below, nine of them). These decomposition mappings, together with similar decomposition mappings for multi-letter Jamos, can be used by processes where one wishes to map between different Jamo sequence representations for
the same text. And that with the same method as for canonical decomposition and composition as is used for normalisation.

With (a certain few) characters already proposed (see L2/07-103) all of the multi-letter Jamos can get auxiliary decomposition mappings similar to canonical decomposition mappings (these not listed in this document).

NOTE, HOWEVER, THAT THE AUXILIARY DECOMPOSITION MAPPINGS CONSIDERED HERE ARE NEVER TO BE USED IN UNICODE NORMALISATIONS.
Not that some of the auxiliary decomposition mappings use Hangul Jamo characters already proposed elsewhere (see L2/07-103).
Proposed additional Hangul Jamo

While I don't think one really should use multi-letter Jamos:

1) many are already encoded,

2) more are on the way (L2/07-103),

3) they are likely to be used, even though Hangul is an alphabetic script and all Hangul texts can be expressed using only Hangul single-letter Jamos,

4) the (few) multi-letter Jamos that occur in decompositions of precomposed Hangul syllables (U+AC00—U+D7A3) will occur in NFD texts formed from texts that use precomposed Hangul syllables,

5) they can be used for ad hoc decomposition and composition to/from single-letter Jamo representation of the text (using the same method as for NFD/NFC).

It is point five here that cannot be fulfilled with the same restrictions as used for canonical decomposition mappings without allocating nine new (not already proposed) multi-letter Hangul Jamos.

In order to get these auxiliary decomposition mappings (not involved in normalisation) to be left-to-right recomposable (after full auxiliary decomposition), and only have two characters in the auxiliary decomposition, nine new characters, Hangul lead consonant Jamos, are hereby proposed:

A97D; <compound> 1100 1109 # NEW proposed HANGUL CHOSEONG KIYEOK-SIOS [NEW PROPOSED CHARACTER; used for the auxiliary decomp. mapping of 3133]
A97E; <compound> 1105 1110 # NEW proposed HANGUL CHOSEONG RIEUL-THIEUTH [NEW PROPOSED CHARACTER; used for the auxiliary decomp. mapping of 313E]
A97F; <compound> 1105 1111 # NEW proposed HANGUL CHOSEONG RIEUL-PHIEUPH [NEW PROPOSED CHARACTER; used for the auxiliary decomp. mapping of 313F]
D7C7; <compound> 1102 1140 # NEW proposed HANGUL CHOSEONG NIEUN-PANSIOS [NEW PROPOSED CHARACTER; used for the auxiliary decomp. mapping of 3168]
D7C8; <compound> A964 1109 # 1105 1100 1109, NEW proposed HANGUL CHOSEONG RIEUL-KIYEOK-SIOS [NEW PROPOSED CHARACTER; used for the auxiliary decomp. mapping of 3169]
D7C9; <compound> A969 1109 # 1105 1107 1109, NEW proposed HANGUL CHOSEONG RIEUL-PIEUP-SIOS [NEW PROPOSED CHARACTER; used for the auxiliary decomp. mapping of 316B]
D7CA; <compound> 1105 1140 # NEW proposed HANGUL CHOSEONG RIEUL-PANSIOS [NEW PROPOSED CHARACTER; used for the auxiliary decomp. mapping of 316C]
D7FC; <compound> 1105 1159 # NEW proposed HANGUL CHOSEONG RIEUL-YEORINHIEUH [NEW PROPOSED CHARACTER; used for the auxiliary decomp. mapping of 316D]
D7FD; <compound> 1106 1140 # NEW proposed HANGUL CHOSEONG MIEUM-PANSIOS [NEW PROPOSED CHARACTER; used for the auxiliary decomp. mapping of 3170]

This proposal does not propose to rearrange the Hangul Jamo characters already proposed (L2/07-103). One may still interfile the code points for the here proposed characters with the ones already proposed, assuming that the code positions for previously proposed Hangul Jamos are not already fixed.