



Document Number: Korea K1496  
 ISO/IEC JTC1/SC2/WG2 N3113

2006-07-14

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 new model for using Hangul Jamo  
**Source:** KIM, Kyongsok; Dae Hyuk Ahn  
**Status:** National Position of Rep. of Korea (KATS)  
**Action:** For consideration by JTC1/SC2/WG2 and UTC  
**Date:** 2006-07-14

*In this document, we propose a new model for using Hangul Jamo and propose to change clause 26.1 of ISO/IEC 10646 accordingly. We also propose to add about 123 characters to the BMP of UCS.*

**Relevant Documents:**

*WG2 N2996 (= K1456) Clause 26.1 of ISO/IEC 10646:2003*  
*WG2 N2994 Request to clarify text on conjoining Jamo clause 26.1*  
*WG2 N3006 Clarification concerning Hangul Syllables 26.1*  
*WG2 N3045 Additions to ISO/IEC 10646:2003*  
*WG2 N3095 (= K1466) Proposed modification to Clause 26.1 of ISO/IEC 10646:2003*  
*WG2 N3104 Resolutions Meeting 48*

**1. According to the resolution M48.31, we propose a new model for using Hangul Jamo.**

*RESOLUTION M48.31 (Model for Hangul Jamos):*

*Republic of Korea is invited to prepare a document on 'a new model for using Hangul Jamos' and send to the convener no later than 2006-07-15, for distribution to and review by WG2 experts.*

As explained below,

- 1) we propose a new model for using Hangul Jamo and propose to change clause 26.1 of ISO/IEC 10646 accordingly.
- 2) we also propose to add about 123 characters to the BMP of ISO/IEC 10646.

## 2. Proposed change to clause 26.1 of ISO/IEC 10646:2003

### 2.1. The Current text of clause 26.1 of ISO/IEC 10646:2003

---- start of the current text of clause 26.1 -----

#### *26.1 Hangeul syllable composition method*

*In rendering, a sequence of Hangeul 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.*

*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 Hangeul syllable composition method.*

*NOTE 1 - Hangeul 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 Hangeul Jamo it should be placed at the end of the sequence, after the Hangeul Jamo character which completes the syllable block.*

---- end of the current text of clause 26.1 -----

## 2.2. A proposed text of clause 26.1

---- start of a proposed text of clause 26.1 -----

### 26.1 Hangul syllable composition method

- a. In rendering, a sequence of Hangul Jamo (from HANGUL JAMO block: U1100 to U11FF and from HANGUL JAMO EXTENDED block: Uxxxx to Uxxxx) is displayed as a series of syllable blocks.
- b. Jamo in Hangul Jamo/Jamo Extended blocks can be classified into three classes: Choseong (= syllable-initial, SI) letters, Jungseong (= syllable-peak, SP) letters, and Jongseong (= syllable-final, SF) letters.
- c. Conceptually, a complete syllable block is composed of a Choseong letter and a Jungseong letter, and optionally a Jongseong letter, where each letter can be either a simple letter (composed of only one simple letter) or a complex letter (composed of two or three simple letters).
- d. Conceptually, an incomplete syllable block is a string of one or more letters which does not constitute a complete syllable. There are four types of incomplete syllable blocks: a Choseong letter alone, a Jungseong letter alone, a Jongseong letter alone, or a Jungseong letter followed by a Jongseong letter. A string of a Choseong letter followed by a Jongseong letter in one block is not considered an incomplete syllable block and, therefore, not allowed.

Examples.

complete syllable blocks: 가, 날

incomplete syllable blocks: ㄷ, ㅏ, ㅑ, 堰

invalid (not allowed): 堰堰堰

The word 'conceptually' in c. and d. above means that we are not talking in terms of code positions; rather, we talk in terms of written letters.

- e. When we use characters from Hangul Jamo and Hangul Jamo Extended blocks, we combine exactly three code positions to represent each complete or incomplete syllable, as shown below: one code position corresponds to each of a Choseong, Jungseong, and Jongseong letter, irrespective of whether each of the letters is simple or complex.
  - 1) a complete syllable block composed of SI and SP letters:
    - SI code position + SP code position + JONGSEONG FILLER (0000 11A7)
  - 2) a complete syllable block composed of SI, SP and SF letters:
    - SI code position + SP code position + SF code position

- 3) an incomplete syllable block composed of an SI letter alone:
    - SI code position + JUNGSEONG FILLER (0000 1160) + JONGSEONG FILLER (0000 11A7)
  - 4) an incomplete syllable block composed of an SP letter alone:
    - CHOSEONG FILLER (0000 115F) + SP code position + JONGSEONG FILLER (0000 11A7)
  - 5) an incomplete syllable block composed of an SF letter alone:
    - CHOSEONG FILLER (0000 115F) + JUNGSEONG HIGH FILLER (0000 11A6) + SF code position
  - 6) an incomplete syllable block composed of SP and SF letters:
    - CHOSEONG FILLER (0000 115F) + SP code position + SF code position
- f. As can be seen, a complex letter cannot be represented as a combination of two or three code positions, where each code position can correspond to a simple or two-complex letter (i.e., a complex letter composed of two simple letters).
- g. The implementation level 3 shall be used for the Hangul syllable composition method.
- h. Usage of code positions from Hangul-related blocks

There are five major categories of blocks containing Hangul letter or syllables.

- 1) Hangul Syllables block (0000 AC00 ~ 0000 D7A3)
- 2) Hangul Jamo/Jamo Extended blocks (0000 1100 ~ 0000 11FF, xxxx xxxx ~ xxxx xxxx)
- 3) Hangul Compatibility Jamo block (0000 3131 ~ 0000 318E)
- 4) Halfwidth and Fullwidth Forms block (0000 FFA0 ~ 0000 FFDC)
- 5) Enclosed CJK Letters and Months block (0000 3100 ~ 0000 321E and 0000 3260 ~ 327E)

Code positions from different categories cannot be combined to make up another syllable block or complex letter.

For example, a code position from Hangul syllables block cannot be combined with code positions from Hangul Jamo/Jamo Extended blocks to make up another syllable block.

Likewise, a code position from Jamo/Jamo Extended blocks cannot be combined with a code position from Hangul Compatibility Jamo block to make up a new syllable block.

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.

---- end of a proposed text of clause 26.1 -----

### **3. Adding Old Hangul complex letters**

We propose to add to ISO/IEC 10646:2003 about 121 Old Hangul complex letters which have been found after the publication of ISO/IEC 10646-1:1993 and, therefore, have not been added to ISO/IEC 10646 yet. We also propose to add two Filler characters in addition: Jungseong High Filler (0000 11A6) and Jongseong Filler (0000 11A7) characters.

We guess that there won't be many Old Hangul complex letters to be found in the future.

### **4. Conclusions**

As explained above,

- 1) we propose a new model for using Hangul Jamo and propose to change clause 26.1 of ISO/IEC 10646 accordingly.;
- 2) we also propose to add about 123 characters to the BMP of ISO/IEC 10646.

By modifying as proposed here, we can disambiguate clause 26.1 and avoid much confusion raised in the past. We are quite sure that methods to process Hangul will become drastically simple and efficient.

Furthermore, when this proposal is accepted, Normalization of Hangul syllables / characters need to be analyzed and modified to accommodate the above changes.

**A list of Old Hangul complex letters with their names will be submitted in a separate document shortly.**

\* \* \*