

Title: Proposal for UTR47
Date: May 9, 2009
From: Ienup Sung <is@sun.com>

L2/09-181

1. Overview

This memo summarizes the notable differences on Hangul processing between the Unicode standard and the RoK NB produced publications (i.e., Korean standards, submissions to WG2 and other standardization organizations, and personal publications from the members of the RoK NB).

For each of the differences, this memo also tries to describe, as proposals, what would be the possible options that we can further discuss and decide on to minimize the differences.

Lastly, this memo also suggests some additional sections that could be added at the UTR47.

By default, this memo does not include any common definitions or specifications among the noted publications unless needed.

2. Conjoining Jamo Behavior and Standard Korean Syllables

The Unicode Standard Version 5.0 (hereinafter, the standard) defines the conjoining jamo behavior at the section 3.12 which also has definitions from D107 to D118. Basically, the standard allows multiple leading consonants (abbreviated each as a L), vowels (abbreviated each as a V), and trailing consonants (abbreviated each as a T) at choseong (첫소리, 초성, 初聲, initial sound, syllable-initial), jungseong (가운뎃소리, 중성, 中聲, middle sound, syllable-peak), and jongseong (끝소리, 종성, 終聲, last sound, syllable-final), respectively. It further clarifies that while in Old Korean, a choseong, a jungseong, and a jongseong can have a sequence of more than one jamo letter at each sound, and in Modern Korean, there is only a single jamo letter at each sound. The standard does not explicitly prohibit the use of conjoining jamo-based Hangul syllables in the everyday Modern Korean writings thereby allowing liberal mixture of jamo-based and pre-composed Hangul syllables in Korean text.

The Korean standard KS X 1026-1:2007 [WG2 N3422R3, L2/08-225] and other publications from the RoK NB (hereinafter, the Korean standard), on the other hand, specify that each of choseong, jungseong, and jongseong can have a single jamo letter defined in the standard. The Korean standard goes further and explicitly specifies that:

4.1 Johap Hangul Letters

Hangul letters specified in Hangul Jamo U+1100 ~ U+11FF, Hangul Jamo Extended-A U+A960 ~ U+A97F, and Hangul Jamo Extended-B U+D7B0 ~ U+D7FF in UCS. These letters are used for composing Old Hangul syllable blocks. ...

5.2 A representation format of Modern Hangul syllable blocks

For representing Modern Hangul syllable blocks, we must use code positions of 11,172 Hangul syllables U+AC00 ~ U+D7A4. ...

5.3 A representation format of Old Hangul syllable blocks

For representing Old Hangul syllable blocks, we must use code positions of Johap Hangul letters in Hangul Jamo U+1100 ~ U+11FF, Hangul Jamo Extended-A U+A960 ~ U+A97F, and Hangul Jamo Extended-B U+D7B0 ~ U+D7FF, adhering to the following three rules. ...

...

In addition, we must adhere to the following rules when representing in Johap Hangul syllable blocks.

- 1) As same as in the rules of representation format for Hangul letters (see 5.1), two or more code positions of simple letters cannot be concatenated to represent a single complex letter.
- an example. ㅏ ㅏ (U+1100 U+1100, incorrect) ⇒ ㅑ (U+1101, correct)
- 2) A Wanseong syllable block cannot be recomposed with Johap Hangul letter(s) to represent another Hangul syllable block.
- an example. ㅑ ㅓ (U+AC00 U+11EB, incorrect) ⇒ ㅑㅓ (U+1100 U+1161 U+11EB, correct)
- 3) A modern syllable block must be represented in Wanseong Hangul syllable block. It is forbidden to represent a modern syllable block in Johap Hangul syllable block.
- an example. ㅏ ㅓ (U+1100 U+1161, incorrect) ⇒ ㅑ (U+AC00, correct)

Here is the summary in a table:

Table 2.1: Differences on Conjoining Jamo Behavior

	<i>The Standard</i>	<i>The Korean Standard</i>
The number of allowed jamo letters in each sound of a jamo-based Hangul syllable:	≥1	1
Allow jamo-based Modern Hangul syllable?	Yes.	No.
Resulting R.E. for Standard Korean Syllables:	(L+ V+ T*) (L* LV V* T*) (L* LVT T*) ☞ Allows conjoining of jamo letter(s) and a pre-composed Hangul syllable to form a new Hangul syllable.	(L V T?) LV LVT ☞ Disallows conjoining of jamo letter(s) and a pre-composed Hangul syllable.

The above differences have implications to several places of the standard, esp., the chapter 3. Conformance, UAX #14 Unicode Line Breaking Algorithm, UAX #29 Unicode Text Segmentation, UTS #10 Unicode Collation Algorithm, and UTS #18 Unicode Regular Expressions.

One thing worth to note at this point is that with the inclusion of the Hangul Jamo Extended-A and Hangul Jamo Extended-B blocks at the standard version 5.2, for all practical purposes, the needs for multiple consonants, vowels, or both in jamo-based Hangul syllables and individual choseong, jungseong, and jongseong representations are, if any, no longer significant. Also, as noted in one of the WG2 submissions from the RoK NB, it appears that any new jamo letters for Old Hangul computing from now on will be also quite rare in numbers if any.

2.1. Proposals

Here are some possible options that we might want to discuss and decide on:

2.1.1. Proposal Option 1:

A) Accept the position specified by the Korean standard in this regard and update the relevant portions of the standard for the next major version to have:

(L V T?) | LV | LVT

B) In the update, do not explicitly specify on the prohibition of jamo-based Modern Hangul syllables so that liberal mixture of jamo-based and pre-composed Hangul syllables can co-exist.

C) In the UTR #47, describe about the transition and provide a mapping table from simple jamo sequences to complex jamo letters shown at the Hangul Jamo U+1100 ~ U+11FF, Hangul Jamo Extended-A U+A960 ~ U+A97F, and Hangul Jamo Extended-B U+D7B0 ~ U+D7FF blocks.

(Note: Considering that the number of jamo-based Hangul syllables that have multiple jamo letters at each sound in real world Korean text is quite rare, this approach will not have big, impact to existing software and Hangul documents.)

2.1.2. Proposal Option 2:

A) Accept the position specified by the Korean standard and update the relevant portions of the standard for the next major version to have:

(L V T?) | LV | LVT

B) In the update, explicitly specify on the prohibition of jamo-based Modern Hangul syllables.

C) In the UTR #47, describe about the transition and provide a mapping table from simple jamo sequences to complex jamo letters shown at the Hangul Jamo U+1100 ~ U+11FF, Hangul Jamo Extended-A U+A960 ~ U+A97F, and Hangul Jamo Extended-B U+D7B0 ~ U+D7FF blocks.

(Note: The difference between the 2.1.1 and the 2.1.2 is on the clause B. This approach will have some impact to existing Hangul documents since people will need to scan, find out any jamo-based Modern Hangul syllables, and convert them to pre-composed Hangul syllables. People will also need to update their software to "forbid" the jamo-based Modern Hangul syllables, figure out an automatic, transparent conversions into pre-composed Hangul syllables, or both.)

2.1.3. Proposal Option 3:

A) Maintain the current position of the standard.

B) In the UTR #47, for all the issues described in the Korean standard in this regard, describe

them and specify alternative resolutions, e.g., implementations should be improved.

While allowing multiple jamo letters in each sound of a Hangul syllable gives some conveniences to theoretical and experimental Hangul studies, clearly that approach has been with some due costs and issues to all of us such as somewhat indeterminate length of Hangul jamo letters in Hangul syllables and corresponding buffer management issues, significant resource requirement to fully support the mixture of jamo and pre-composed Hangul syllables, and it also contributed to unresolved issues requiring interim solutions for Hangul collation noted in the UTS #10.

With the new Hangul Jamo Extended-A and -B blocks and by limiting the number of jamo letters to one (or, in case of jongseong, one or zero), and thereby changing the model of conjoining jamo behavior, we can bring more certainty and simplifications in Hangul computing and software.

3. Hangul Syllable Boundary Determination

Due to the differences described at the previous section, the standard and the Korean standard have differences in the syllable break opportunities as outlined in the following table. Note that the common rows are marked with bold-italic typeface style and a highlighted color:

Table 3.1: Differences on Hangul Syllable No-Break Rules

<i>Do Not Break Between</i>		<i>Examples</i>
<i>L</i>	L, <i>V</i> , or precomposed Hangul syllable	L × L <i>L × V</i> L × LV L × LVT
<i>V</i> or LV	V or <i>T</i>	V × V <i>V × T</i> LV × V LV × T
T or LVT	T	T × T LVT × T
<i>Jamo or precomposed Hangul syllable</i>	<i>Combining marks</i>	<i>L × M</i> <i>V × M</i> <i>T × M</i> <i>LV × M</i> <i>LVT × M</i>

3.1. Proposals

3.1.1. Proposal Option 1:

A) Accept the following as the no-break rules for Hangul Syllables and update the relevant portions of the standard for the next major version:

Table 3.2: Proposed Hangul Syllable No-Break Rules

<i>Do Not Break Between</i>		<i>Examples</i>
<i>L</i>	<i>V</i>	<i>L × V</i>
<i>V</i>	<i>T</i>	<i>V × T</i>
<i>Jamo or precomposed Hangul syllable</i>	<i>Combining marks</i>	<i>L × M</i> <i>V × M</i> <i>T × M</i> <i>LV × M</i> <i>LVI × M</i>

B) Along with the clause C of the sections 2.1.1 and 2.1.2, describe the transition in this regard also at the UTR #47.

3.1.2. Proposal Option 2:

A) Maintain the current position of the standard.

B) In the UTR #47, for all the issues described in the Korean standard in this regard, describe them and, optionally, specify alternative resolutions.

4. Unicode Normalization Forms and Hangul Characters

Concerning on the Unicode Normalization Forms with Hangul characters, basically, we have three issues that could be seen as different between the standard and the Korean standard. The following three subsections describe on the differences and propose possible resolutions to discuss and decide on.

4.1. Hangul Compatibility and Halfwidth Jamo Letters

Due to that the compatibility decomposition transforms Hangul Compatibility Jamo letters at U+3131 ~ U+318E and half-width Hangul jamo letters at U+FFA0 ~ U+FFDC of Halfwidth and Fullwidth Forms block ultimately into Hangul Jamo letters at U+1100 ~ U+11FF, without any intervention, transformed jamo letters could conjoin with adjacent jamo letters in display and also transform further into pre-composed Hangul syllables in the NFKC form.

The standard specifies that they, Compatibility and Halfwidth Jamo letters, are "spacing, nonconjoining Hangul consonant and vowel (jamo) elements" and the characters "are provided solely for compatibility with the KS X 1001:1998 standard." It further specifies that, when necessary, in the NFKD and the NFKC forms, separation of such jamo letters can be achieved by using either U+200B ZERO WIDTH SPACE or U+2060 WORD JOINER as shown in the following (which is the *Table 12-11. Separating Jamo Characters* at the standard):

Table 4.1: Separating Jamo Characters (The Unicode Standard)

Original	NFKD	NFKC	Display
ㄱ ㅏ 3131 314F	ㄱ ㅏ 1100 1161	가 AC00	가
ㄱ ZW SP ㅏ 3131 200B 314F	ㄱ ZW SP ㅏ 1100 200B 1161	ㄱ ZW SP ㅏ 1100 200B 1161	가

For the NFKD and the NFKC forms, the Korean standard recommends to insert filler characters so that Hangul Jamo letters originated from the Hangul Compatibility Jamo and the Halfwidth Hangul jamo letters can be processed as "correct Hangul syllable" characters. Here is the table showing an example on how that should be done from the KS X 1026-1:2007:

Table 4.2: An Example on Hangul Compatibility Jamo Letters to Standard Korean Syllables (The Korean Standard)

Original	NFKD	NFKC
ㄱ ㅏ	ㄱ HJ F HC F ㅏ	ㄱ HJ F HC F ㅏ
U+3131 U+314F	U+1100 U+1160 U+115F U+1161	U+1100 U+1160 U+115F U+1161

This is due to that the RoK NB sees, as shown in the sections 4.1, 4.2, 4.3, and 6.2 of the KS X 1026-1:2007:

- The Hangul Jamo, Hangul Jamo Extended-A, and Hangul Jamo Extended-B letters are used only for composing Old Hangul syllables and are not used alone.
- The compatibility and half-width jamo letters should not be used to form syllables; they are to represent full-width Hangul consonants and vowels individually and for the compatibility with the 7-bit Hangul code in KS X 1001, respectively.

The differences can be summarized into the following table:

Table 4.3: Differences on How to Handle Hangul Compatibility & Halfwidth Jamo Letters with the NFKD and the NFKC Forms

	<i>The Standard</i>	<i>The Korean Standard</i>
Should jamo letters from Hangul Compatibility Jamo and Halfwidth be treated as Standard Korean Syllables?	No.	Yes.
How to separate jamos from Hangul Compatibility Jamo and Halfwidth and Fullwidth Forms?	Insert U+200B ZERO WIDTH SPACE or U+2060 WORD JOINER where the separation is needed.	Insert proper filler characters to convert all occurrences into Standard Korean Syllables.

4.1.1. Proposal Option 1:

A) Maintain the current position of the standard.

B) Describe the differences at the UTR #47 and give an option to people saying that, as a pre-processing, people may can insert U+115F HANGUL CHOSEONG FILLER, U+1160 HANGUL JUNGSEONG FILLER, or both instead of the U+200B ZERO WIDTH SPACE or U+2060 WORD JOINER to prepare to transform such jamo letters into Standard Korean Syllables.

4.1.2. Proposal Option 2:

A) Accept the position specified in the Korean standard on this and update the relevant portions of the standard for the next major version. This will be described as a pre-processing in the standard.

B) In the UTR #47, describe the transition and also a pre-processing procedure on how to convert any of the existing Hangul text with the current scheme into the new way.

4.2. Parenthesized and Circled Hangul Elements

In the same context as described in the previous section, the RoK NB sees that parenthesized and circled Hangul jamo elements being transformed into Hangul Jamo block letters through the compatibility decomposition should be treated as Standard Hangul Syllables. The Korean standard, thus, recommends that when transforming parenthesized and circled Hangul jamo characters to the NFKD or the NFKC forms, append the U+1160 HANGUL JUNGSEONG FILLER character to each of the occurrences so that they will be Standard Korean Syllables. The following example from the Korean standard explains that:

Table 4.4: An Example of Circled Hangul Element to Standard Korean Syllable
(The Korean Standard)

Original	NFD	NFC	NFKD	NFKC
ㄱ	ㄱ	ㄱ	ㄱ HJ F	ㄱ HJ F
U+3260	U+3260	U+3260	U+1100 U+1160	U+1100 U+1160

Naturally, the above recommendation is only for Parenthesized Hangul elements from U+3200 PARENTHESESIZED HANGUL KIYEOK (ㄱ) to U+230D PARENTHESESIZED HANGUL HIEUH (ㅎ) and Circled Hangul elements from U+3260 CIRCLED HANGUL KIYEOK (ㄱ) to U+326D CIRCLED HANGUL HIEUH (ㅎ).

The standard simply specifies the compatibility decomposition mappings and some other minor descriptions that are not relevant to the subject matter.

4.2.1. Proposal Option 1:

A) Maintain the current position of the standard.

B) Describe the differences at the UTR #47 and give an option to people saying that, as a pre-processing, people may can replace each of the Parenthesized Hangul elements with a sequence of an opening parenthesis, a corresponding Hangul Jamo letter, a U+1160 HANGUL JUNGSEONG FILLER, and a closing parenthesis. Similarly, also describe that, as a pre-processing, people may can replace each of the Circled Hangul elements with a corresponding Hangul Jamo letter followed by a U+1160 HANGUL JUNGSEONG FILLER.

C) Provide a mapping table for the replacement procedure as a Reference at the UTR #47.

4.2.2. Proposal Option 2:

A) Accept the position specified in the Korean standard on this subject and update the relevant portions of the standard for the next major version. This will be described as a pre-processing in the standard.

B) In the UTR #47, describe the transition and also the replacement procedure on the Parenthesized and Circled Hangul elements as a pre-processing.

C) Provide a mapping table for the replacement procedure as a Reference at the UTR #47.

4.3. Old Hangul Jamo-based Syllables

The standard specifies that the jamo letters can conjoin with pre-composed Hangul syllables. As a result, during the canonical composition, any adjacent jamo letters that can be transformed into a Modern Hangul syllable, they will be. This includes cases where the jongseong of the corresponding syllable has Old Hangul jamo as shown in the following example from the Korean standard:

Table 4.5: Old Hangul Syllables in Different Representations in Storage
(and Also on The Majority of The Current Displays)

Original	NFC	NFKC
갇	가 △	가 △
U+1100 U+1161 U+11EB	U+AC00 U+11EB	U+AC00 U+11EB

As described in the section 2, the RoK NB sees it differently that each syllable should be either a pre-composed Modern Hangul syllable or a conjoining jamo-based Hangul syllable, in particular, for Old Hangul only. In other words, the RoK NB recommends that the conjoining jamo-based Hangul syllables must retain the same code points when there is Old Hangul jamo letters in the syllable regardless of the Unicode Normalization forms:

Table 4.6: Recommended Old Hangul Representations by The Korean Standard

Original	NFC	NFKC
꺄	꺄	꺄
U+1100 U+1161 U+11EB	U+1100 U+1161 U+11EB	U+1100 U+1161 U+11EB

The difference can be summarized into the following table:

Table 4.7: Differences on Old Hangul Syllables

	<i>The Standard</i>	<i>The Korean Standard</i>
Allow conjoining of pre-composed Modern Hangul syllables with conjoining Old Hangul jamo letters?	Yes.	No.

4.3.1. Proposal Option 1:

A) Maintain the current position of the standard.

B) Describe the difference at the UTR #47 and give an option to people saying that, as a pre-processing or during input processing (automatically or manually), prior to the NFC or the NFKC processing, for all conjoining jamo sequences, when there is a consecutive sequence of L V T where the L and the V are not fillers and the T is one of the Old Hangul jamo letters, then, U+200D ZERO WIDTH JOINER can be inserted between the L and the V so that during the canonical composition, the conjoining jamo sequences will retain the same code points.

C) Provide a list of the L V T sequences that may need the insertion of the U+200D ZERO WIDTH JOINER to retain the same code points after the NFC or the NFKC processes as a Reference at the UTR #47.

D) Describe that, optionally, after the Unicode Normalization, the U+200D ZERO WIDTH JOINER may also be removed from the Old Hangul jamo sequences.

4.2.2. Proposal Option 2:

A) Accept the position specified in the Korean standard on this subject and update the relevant portions of the standard for the next major version. This will be described as pre- and optional post-processing mechanisms in the standard with the pre- and the post-processing mechanisms described in the section 4.3.1.

B) In the UTR #47, describe the transition and also the pre- and the optional post-processing mechanisms as additional procedures.

C) Provide a list of the L V T sequences that may need the insertion of the U+200D ZERO WIDTH JOINER to retain the same code points after the NFC or the NFKC processes as a Reference at the UTR #47.

5. Hangul Collation

There are mainly two differences between the standard and the Korean standard for Hangul collation:

- Collation weight definition between the UTS #10 DUCET and the Korean Standard.
- On how to form sort keys and compare.

The following two subsections describe the differences in detail and propose possible resolutions.

5.1. Differences on Collation Weight Definitions

While the DUCET is a default collation element table and people can provide their own through a customization, the following differences are noted from the current DUCET (<http://www.unicode.org/Public/UCA/latest/allkeys.txt>) and the order definitions at the Korean standard:

Table 5.1: Order Differences in The Same Group—Hangul Jamo Choseong Letters, Left-to-right and Top-to-Bottom Order Example

<i>The Standard</i>	ㄱ (U+1100), ㅋ (U+1101), ㆁ (U+1102), ㄷ (U+1103), ㄹ (U+1104), ...
<i>The Korean Standard</i>	ㄱ (U+1100), ㅋ (U+1101), ㆁ (U+1102), <u>ㄷ (U+1113), ㄴ (U+1114), ㄷ (U+1115),</u> <u>ㄹ (U+1116), ㄷ (U+1103), ...</u>

As shown in the above example, there are some significant differences on the order of Hangul jamos between the two standards. As also shown in the table at below, there are differences on how the orders should be defined among Hangul characters from different blocks:

Table 5.2: Order Differences Among Groups—KIYEOK, Top-to-bottom Order Example

<i>The Standard</i>	<i>The Korean Standard</i>
U+1100 HANGUL CHOSEONG KIYEOK ㄱ	U+1100 HANGUL CHOSEONG KIYEOK ㄱ
U+3131 HANGUL LETTER KIYEOK ㅋ	<u>U+11A8 HANGUL JONGSEONG KIYEOK ㆁ</u>
U+3200 PARENTHESIZED HANGUL KIYEOK (ㄱ)	<u>U+FFA1 HALFWIDTH HANGUL LETTER KIYEOK ㄱ</u>
U+3260 CIRCLED HANGUL KIYEOK ㄱ	<u>U+3131 HANGUL LETTER KIYEOK ㅋ</u>
U+FFA1 HALFWIDTH HANGUL LETTER KIYEOK ㄱ	<u>U+3200 PARENTHESIZED HANGUL KIYEOK (ㄱ)</u>
U+320E PARENTHESIZED HANGUL KIYEOK A (ㄱ)	<u>U+3260 CIRCLED HANGUL KIYEOK ㄱ</u>
U+326E CIRCLED HANGUL KIYEOK A (ㄱ)	<u>U+AC00 HANGUL SYLLABLE GA ㄱ</u>
U+AC00 HANGUL SYLLABLE GA ㄱ	<u>U+320E PARENTHESIZED HANGUL KIYEOK A (ㄱ)</u>
U+11A8 HANGUL JONGSEONG KIYEOK ㆁ	<u>U+326E CIRCLED HANGUL KIYEOK A (ㄱ)</u>

Lastly, there is also a difference in defining the order between Choseong, Jungseong, and Jongsung groups as shown in the following table. Please also note that, as we all know, this also has an implication to Hangul syllable to Hangul syllable collation order comparison:

Table 5.3: Order Differences Among Choseong, Jungseong, and Jongseong Letters, Top-to-bottom Order

<i>The Standard</i>	<i>The Korean Standard</i>
Choseong Jungseong Jongseong	Choseong <u>Jongseong</u> <u>Jungseong</u>

5.1.1. Proposal Option 1:

A) Maintain the current position of the standard.

B) In the UTR #47, describe that the DUCET defined weights for Hangul characters are not widely accepted by Korean society and mention that, if needed, the collation element table can be customized. Explain how to customize in accordance with the Korean standard.

C) In the UTR #47, provide a customized collation element table for Korean characters.

5.1.2. Proposal Option 2:

A) Update the DUCET in accordance with the Korean standard and update the relevant portions of the standard for the next major version.

B) In the UTR #47, describe on the transition of the DUCET definition and explain the collation order and the principles used in the collation weight definition.

5.2. Difference on How to Form Sort Keys and Compare

While the standard recognizes Hangul sorting problems and provides three interim solutions at the section 7.1.4.1 Hangul Trailing Weights of the UTS #10, fundamentally, it is based on *per-character* sort key formation and comparison. On the other hand, the Korean standard treats all Hangul characters as, in a sense, *syllabic entities* thereby resolving the sorting problem. Another thing to note in the Korean standard is that there is no need for the "contractions" since there is only an L, a V, and an optional T letter in a Hangul syllable (or "syllabic entity").

The differences can be summarized into the following:

Table 5.4: Differences on Sort Key Formation and Comparison

	<i>The Standard</i>	<i>The Korean Standard</i>
Unit of sort key formation and comparison:	A jamo.	A syllable (or, syllabic entity).
Sorting is done by using:	Collation weight definition data <i>without</i> extra algorithmic help.	Collation weight definition data <i>with</i> extra algorithmic help.
Can it handle multi-letter Choseong, Jungseong, and Jongseong?	No.	N/A.
Can Hangul syllables be sorted correctly in mixed script text?	No.	Yes.

5.2.1. Proposal Option 1:

- A) Update the DUCET as described in the section 5.1.2.
- B) Update the UTS #10 for the next major version of the standard so that in the section 4.2 Produce Key, the algorithm will recognize Hangul syllables without T and append a pre-defined weight W_{T_0} that is $W_{T_{last}} < W_{T_0} < W_{V_{first}}$ for each of such Hangul syllables where the $W_{T_{last}}$ is the weight for the last jongseong and the $W_{V_{first}}$ is the weight for the first jungseong in the collation element table.
- C) Update or delete the section 7.1.4.1 Hangul Trailing Weights of the UTS #10 for the next major version of the standard.
- D) In the UTR #47, describe on the transition.

5.2.2. Proposal Option 2:

- A) Maintain the current position of the standard.
- B) In the section 7.1.4.1 Hangul Trailing Weights of the UTS #10, add another interim solution described in the previous section.
- C) In the UTR #47, further clarify the issues and provide an algorithm that can be used with the customized collation element table.

6. Hangul Punctuation and Tone Marks

I would like to propose to add a section in the UTR #47 that is dedicated to the explanations on what are the punctuation and tone marks used in Hangul text and also what would be the proper Unicode characters for both horizontal and vertical writings.

7. Code Conversions

There are several code conversion related *gotchas* when you develop the code conversions between the standard and other traditional/legacy Korean character sets and encodings that we do not have a place to add info to. I would like to also propose to add a section at the UTR #47 that will deal with such issues and clarify on how such boundary code conversion cases should be handled.

END_OF_MEMO.