

## Request to provide explicit rationales for exclusion of symbols in the KPS 9566 standard

Eduardo Marín Silva

27/01/2018

**Introduction.** In a previous document (<https://www.unicode.org/L2/L2018/18004-compat-dprk.pdf>) I proposed to reconsider compatibility symbols used by the KPS standard used in North Korea. I made that proposal since there was no clear rationale for their exclusion (apart from possibly the HAMMER AND SICKLE AND BRUSH). During UTC 154, my proposal was discussed and denied on the grounds that “Unless there is actual demand to encode these characters, beyond simply being present in a standard that is effectively unsupported outside of DPRK, the characters are not deemed suitable for encoding.”. Here I explain how such an answer is not in line with the consortium’s stated goal.

**Refuting the notion that they are not suitable.** The fact that the committee did not provide what I consider “valid” reasons for exclusion, such as unification or being trademarked, confirms my suspicion that they were originally arbitrarily excluded. The characters in question are still used except for the enclosed postal mark which was removed (further discussion on that sign deserves its own document), the standard was updated three times, in 2000, 2003 and 2011 respectively, which shows that it is still being used. The fact that there is no full roundtrip compatibility with Unicode would make it unique amongst all national standards.

All standard should be equal under the committee’s view. There is no precedent to exclude standards based on popularity. It could be argued that reaching roundtrip compatibility with KPS would somehow imply allegiance to the regime, but that would be a fallacy. I for instance detest the regime of North Korea and acknowledge its immense human rights problem, however if peace is ever to be reached between the Koreas, they should be treated as equals.

**Status quo.** Unicode was conceived to ease information interchange amongst the world, but right now, the two Koreas cannot interchange documents without some information being lost.





For now, North Korean users are forced to use PUA codepoints to represent their missing symbols as Jaemin clearly showed (<https://www.unicode.org/L2/L2018/18011-info-kps9566-2011.pdf>), use of PUA is supposed to be a last resort but this simply does not apply to the symbols proposed.

The consortium has been known to change its mind, for instance, encoding separate chillus for Malayalam or disunifying the SIDDHAM LETTER THREE-CIRCLE ALTERNATE I, were considered inadmissible until a need was shown. It is never too late to reconsider, even with the amount of time that has passed.



**Hammer and sickle and brush.** I think that speculating about possible controversies is not helpful, as I stated before, there are many symbols that are similar in function, the only time a great deal of care was put into encoding symbols not particular to a script were the svastis, however the inclusion of the INVERTED PENTAGRAM did not cause much fuss. We must also take into account that the trigrams present in the South Korean flag are already encoded. We won’t know the true opponents of this symbols until they are presented up for a vote.

**Map symbols.** The only change made here, was to revise the name of the first character, in order to be more like OUTLINED BLACK STAR. I also learned that the last three are used to indicate slopes in the way, and so are not necessarily for use in battleground maps.



Glyph	KPS code	Proposed codepoint	Proposed name
	A2F1	U+1F7D9	UP POINTING OUTLINED BLACK TRIANGLE
	A2F2	U+1F7DA	UP POINTING BLACK TRIANGLE WITH HORIZONTAL FILL
	A2F3	U+1F7DB	UP POINTING BLACK TRIANGLE WITH UPPER LEFT TO LOWER RIGHT FILL
	A2F4	U+1F7DC	UP POINTING BLACK TRIANGLE WITH UPPER RIGHT TO LOWER LEFT FILL



**Circled white up pointing index.** I still maintain the same rationale for inclusion and preferred codepoint at 1F90C.

**Presentation forms of punctuation.** I do not propose those anymore, since according to Jaemin they are not mapped to PUA, so there is apparently a better mechanism to handle such characters.

**Double punctuation.** I still maintain the same rationale for inclusion, either as separate characters with decompositions or as named character sequences.

Glyph	KPS code	Proposed codepoint	Proposed name	Decomposition or named sequence
	A1DC	U+2E50	FULL STOP RIGHT PARENTHESIS	≈ 002E . 0029 )
	A1DD	U+2E51	FULL STOP RIGHT DOUBLE ANGLE BRACKET	≈ 002E . 300B »

**Vulgar fractions.** I maintain the same rationale for their inclusion as standardized variation sequences.

- 00BC FE00; horizontal stroke form; # VULGAR FRACTION ONE QUARTER
- 00BD FE00; horizontal stroke form; # VULGAR FRACTION ONE HALF
- 00BE FE00; horizontal stroke form; # VULGAR FRACTION THREE QUARTERS
- 2153 FE00; horizontal stroke form; # VULGAR FRACTION ONE THIRD
- 2153 FE00; horizontal stroke form; # VULGAR FRACTION TWO THIRDS

**Emphazised Hangul syllables.** After learning that modern font technology can add the emphasis automatically and fairly easily, I see no reason for encoding them in any way. Except if we want to handle the emphasis at the CLDR level the named character sequences may be useful.