_____

Author: KIM, Kyongsok
Date: 2012.06.09.
Status: NB position of R.O.Korea
Subject: Comments RE: Providing Total Strokes in Submissions
Relevant document(s): IRG N1810 (IRG #37 Res. M37.13), IRG N1789, IRG N1790

1. Related resolution documents

1.1 IRG N1810 (IRG #37 Resolution M37.13)

Resolution IRG M37.13: Providing Total Strokes in Submissions

Unanimous

In response to WG2 request(IRGN 1789, IRGN 1790) for an additional total stroke count requirement, some IRG members need more time to consider the implication for future submissions although some members have considered it acceptable. The IRG further request its member bodies to give their feedback for discussion in IRG #38.

1.2 WG2 res. M58.27 (2011.06.10., HELsinki, WG2 N4104=IRG N1789)

RESOLUTION M58.27 (P&P document – adoption and additions):

Unanimous

WG2 adopts document N3902 as the updated principles and procedures. Further, WG2 accepts the additions proposed under item A, item B (with slight modification as recorded in the minutes) and item C, in document N4049, and instructs its ad hoc on Principles and Procedures to update the P&P document (as document N4102) for adoption at WG2 meeting 59. IRG is requested to review and provide feedback on item D in document N4049.

1.3 Item D. Stroke Counts for Ideographs in WG2 N4049=IRG N1790

D. Stroke Counts for Ideographs

Add the following new question between current questions 12 and 13 in section C of the proposal summary form:

1x. Does the proposal contain any Han Ideographs? _____
If YES, is the total number of strokes (including the radical) for each glyph to be shown in the UCS charts specified? _____

Rationale:
Over the course of checking data in the Unihan database, a number of errors in the kTotalStrokes field were identified. This field contains the total number of

strokes in the character (including the radical). Its value is for the character as drawn in the UCS charts. *The total number of strokes is frequently used for sorting ideographs,* and so getting these values correct is important. In order to expedite keeping this data current, the UTC has requested that the IRG have total stroke counts included *in future submissions*. Since some of the ideograph submissions are dealt with directly in WG2, a similar request is being made here to submitters to WG2 also.

2. Analysis

1) As pointed out in WG2 N4049 (=RIG N1790), the total number of strokes is frequently used in sorting/searching CJK characters.

2) However, the total number of strokes including the radical may be different depending on the countries.

3) It seems that, since only one total number of strokes will be kept in the UniHan database, the number will be the Chinese-style total number of strokes.
   - In other words, the total number of strokes provided by countries other than China need to be "verified" by China.

4) Therefore, even if the total number of strokes is kept in the UniHan database, applications to be used in countries/regions other than in China might use a total number of strokes which is different from that in Unihan database.

5) Question: R.O.Korea wonders if the request (i.e. *Providing Total Strokes in Submissions*) implies that IRG reviews and decides on the total number of strokes including the radical.
   - If the answer is yes, review processing at IRG will be slowed down.

6) Once the radical and the number of strokes of the remaining part the CJK characters are provided as is done currently, one can calculate the total number of strokes by simply adding the two numbers:

   a) "Chinese-style" number of strokes of the radical in a CJK char; and
   b) the number of strokes of the remaining part the CJK characters, which is currently provided.

7) The proposal does not suggest to put the total number of strokes in ISO/IEC 10646.

3. A conclusion

   Based on the above analysis, it might be better to decide on the Chinese-style total number of strokes by the direct cooperation between China and Unicode rather than by going through IRG and/or WG2.

* * *