ISO/IEC JTC1/SC2/WG2/IRG N1614
Doc. #: Korea JTC1/SC2 K1833

Korea JTC1/SC2, Committee on Character Codes

KIM, Kyongsok (GIM, Gyeongseog), Chairperson of Korea JTC1/SC2

email: gimgs0 AT gmail DOT com, phone: +82-51-510-2292
address: Division of Computer Science and Engineering,
Pusan National University
98 Busandaehagro Geumjeonggu, BUSAN 609-735, Rep. of KOREA

Author: KIM, Kyongsok (GIM, Gyeongseog), Chairperson of Korea JTC1/SC2
Date: 2009.10.28
Status: National Body Position, Rep. of KOREA
Subject: Korea's comments on SC2 N4079, CD Ballot on ISO/IEC 10646, 2ed.

* Only the portions of Korea JTC1/SC2 K1822 (= WG2 N3686, 2009.09.24) relevant to IRG are shown below. (2009.10.28)

[Technical comments]

T2. p.40, 23.2, third bullet

3rd field: Hanzi G sources(G0-hhhh), (G1-hhhh), (G3-hhhh), (G5-hhhh), (G7-hhhh), (GS-hhhh), (G8-hhhh), (G9-hhhh), (GE-hhhh), (G_4K), (G_BK), (G_BKddddd), (G_CH), (G_CY), (G_CYYddddd), (G_CHddddd), (G_FZ), (G_FZddddd), (G_GHddddd), (G_GFHZBddd), (G_GJZddddd), (G_HC), (G_HCddddd), (G_HZ), (G_HZddddd), (G_IDCddd), (G_XCddddd), (G_ZFYddddd), (G_ZJWddddd), or (KXdddd.dd)

[current text] KXdddd.dd

--->
[proposed text] GKXdddd.dd (or G_Kdddd.dd or G_Xdddd.dd or G_KXdddd.dd)

Rationale: Since all G sources except KX start with 'G', we propose to change KXdddd.dd to GKXdddd.dd (or G_Kdddd.dd or G_Xdddd.dd, or G_KXdddd.dd) so that all G sources start with 'G'. Furthermore, considering that all Hanja K sources starts with 'K' and and Hanja KP sources starts with 'KP', 'G_K...' or 'G_X...' look much better than

'KX...'

T3. p.41, 23.3, third para.

The code chart for the CJK UNIFIED IDEOGRAPHS block (4E00-9FFF) uses a fixed column format (i.e. source references from a given source always appear in the same column) while the code charts for the other CJK Unified blocks show graphic symbols per the following order of appearance: G, T, J, K, V, KP, H, U, and M

[current text] G, T, J, K, V, KP, H, U, and M

--->
[proposed text] G, T, J, K, KP, V, H, M, and U

Rationale:
1) KP precedes V (e.g., p. 388, U340C)
2) In CJKU main, M (in the second column) precedes U in the third column).
Note. In CJKU Extensions, there is no Hanja char having both M and U source references.

T4. p. 42, 23.3.2

23.3.2 Source reference presentation for CJK UNIFIED IDEOGRAPHS EXTENSION A
The following figure shows the presentation for the CJK UNIFIED IDEOGRAPH EXTENSION A block. Up to four sources per characters are represented in a single row. If more than four sources exist, an additional row is used.

[current text] ... four sources ...

--->
[proposed text] ... three sources ... [occurs twice]

T5. p. 43
1) There is no explanation as to which country will provide font for CJK Compatibility characters shared by more than one country. We suggest to discuss and to add an explanation to the Standard.
For your information, the sharing status is as follows:
KJ 6, KPJ 1, KH 1, PT 49, PTH 1, TH 4, 62 in total (out of 1,000 CJKC

chars)

2) Rep. of Korea will provide the font for U0F900 ~ U0FA0B and requests that the font be used for printing U0F900 ~ U0FA0B.

T16. CJKU_SR.txt and CJKC_SR.txt

[current text]
- In CJKU_SR.txt, UTC is used
- In CJKC_SR.txt, U0- is used

--->
[proposed change]
- We propose to use consistently either UTC or U0- in both CJKU_SR.txt and CJKC_SR.txt.

T17. CJKC_SR.txt

[current text]

As an example, there is an entry where a UCS code position and a U0 source ref. value are the same, as shown below:

0FA0C;05140;;;;;U0-FA0C;

--->
[proposed change]
- We wonder what useful information a user/reader can get from "U0-FA0C" in this example. The code positions are the same.
- There are 21 more code positions having the same situation.
- Unless U0-xxxx can give some useful information, we suggest to delete those 22 U0- entries in CJKC_SR.txt

T19. p. 1218, left column

T19.1 We request to change as shown below:

T19.2 We request to change the following line in CJKC_SR.txt as shown below:

```
(current) 0F9B8;096B8;;;;KO-6766;;
--->
(new)     0F9B8;096B7;;;;KO-6766;;
```

== Rationale (Information supporting our request):

a) By checking the glyphs in 2ed CD, we can see that UF9B8 should be mapped to U96B7, not to U96B8.

b) Furthermore, duplicate Hanja characters are included in KS X 1001 (K0), but not in KS X 1002 (K1).
  - Therefore, any compatibility Hanja characters (whose source is K0, including UF9B8) must be mapped to a K0 Hanja (in this case, U96B7), but not to K1 Hanja (in this case, U96B8).

c) In CJKU_SR.txt, we know that U96B7 is a K0 Hanja and U96B8 is a K1 Hanja.

```
096B7;GE-443F;T3-5349;J0-4E6C;K0-564B;;;KP0-FDB7;;
096B8;G1-4125;T1-7622;J0-7031;K1-5E68;;;KP1-83A8;;
```

d) Mapping info RE: duplicate Hanja in KS X 1001 (and comp. Hanja in UCS)
  - source: Korea JTC1/SC2 documents K1645 and K1646
    (= SC2/WG2 N3420 and 3421, respectively).

ro-co	KSX1001(=EUC-KR)	UCS		ro-co	KSX1001(=EUC-KR)	UCS
71-70	0x6766 (=E7E6)	U+F9B8 隸 예	=	54-43	0x564B (=D6CB)	U+96B7 隸 례

e) Exact glyphs of two Hanja characters in KS C 5601 are shown below:
  - 71-70 0x6766 (=E7E6) and 54-43 0x564B (=D6CB)

KS C 5601:1987 (= KS X 1001), International Register 149 (http://www.itscj.ipsj.or.jp/ISO-IR/149.pdf)

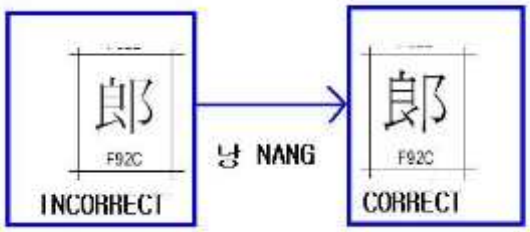
71-70 0x6766 (=E7E6)
예 (ye)

隸

54-43 0x564B (=D6CB)
례 (rye)

T20. pp. 1213 and 1215; CJKC_SR.txt

T20.1 p 1213: We request to change the glyph for U+F92C as shown below:
- We need to add one more stroke (i.e., The number of strokes need to be changed from 10 to 11).



T20.2 p. 1215; We request to change two lines as shown below:
T20.2.1) Change 90CE to 90DE
T20.2.2) Change glyphs of two Hanja characters from 10 strokes to 11 strokes.

T20.3 We request to change one line in CJKC_SR.txt file as shown below:

```
0F92C;090CE;;;;K0-522B;;
---->
0F92C;090DE;;;;K0-522B;;
```

== Rationale (Information supporting out request):

a) Mapping info RE: duplicate Hanja in KS X 1001 (and comp. Hanja in UCS)
- source: Korea JTC1/SC2 documents K1645 and K1646
(= SC2/WG2 N3420 and 3421, respectively).

- 낭 NANG K0 0x522B, (50-11: row-col), 0xD2AB, U+F92C
랑 RANG K0 0x554D, (53-45: row-col), 0xD5CD, U+90DE

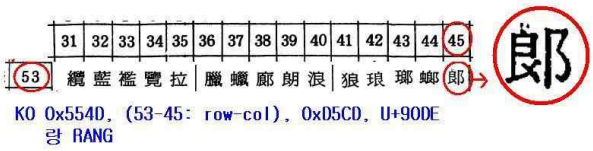
b) Exact glyphs of two Hanja characters in KS C 5601 are shown below:
- 50-11 0x522B (=D2AB) and 53-45 0x554D (=D5CD)
- As we can see, their glyphs are exactly the same.
- source: KS C 5601-1987 (<-- International Register 149)
(http://www.itscj.ipsj.or.jp/ISO-IR/149.pdf)

- The number of strokes for these two characters is 8 + 3 = 11,
not 7 + 3 = 10.
(Note: The number of strokes could be 10/9 instead of 11/10. In this document, we will use 11/10).

http://www.itscj.ipsj.or.jp/ISO-IR/149.pdf

TYPE: Multiple-byte Graphic Character Set	REGISTRATION NUMBER: 149 DATE OF REGISTRATION: 1st Oct.1988
ORIGIN Korean Standard KS C 5601-1987	

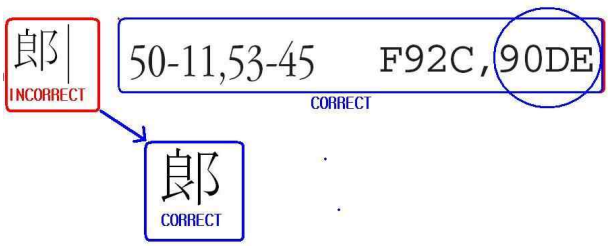
KO 0x522B, (50-11: row-col), 0xD2AB, U+F92C
낭 NANG



KO 0x554D, (53-45: row-col), 0xD5CD, U+90DE
랑 RANG

c) If the glyph of U+F92C (0x522B, 50-11, 낭 Nang) WERE correct (10 strokes),
- since the glyph of U+F92C (0x522B) is different from the glyph of 랑 Rang (0x554D, 53-45, 11 strokes), U+F92C (0x522B) SHOULD NOT HAVE BEEN encoded as compatibility Hanja.
- Instead, we could simply fill the "currently empty" K column for U+90CE with "KO-522B/KO-5051".
- Therefore, we can conclude that the glyphs of "낭 Nang (0x522B, 50-11)" and "랑 Rang (0x554D, 53-45)" in KS C 5601 are exactly the same and the number of their strokes is 11.

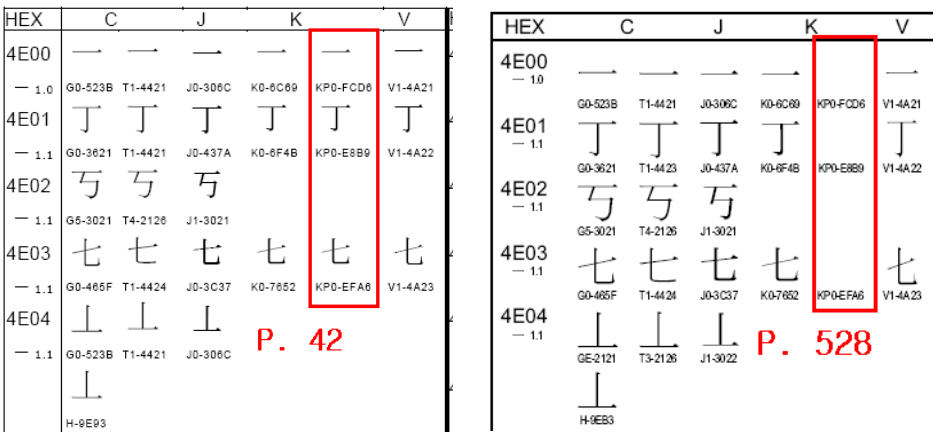
d) (This is informational)
- In Ken Lunde's book, CJKV Information Processing, UF92C is correctly mapped to U90DE, which is another evidence supporting our request of change.
- However, the glyph is incorrect. We need to add one more stroke (i.e., The number of strokes need to be changed from 10 to 11).
- He promised that he would correct the glyphs.



＊ ＊ ＊

[Editorial comments]

E1. pp. 42, 23.3.1 and 528.
- Due to font problems, KP columns are different on pages 42 (Fig. 2) and 528.
- We expect that fonts will solve this discrepancy.



＊ ＊ ＊