

Minutes of the Editorial Group Ad Hoc Discussion on Member submitted documents (N2664, N973, N1006, N972, N1008r)

Discussion Item 1.

The Editorial Group studied the different glyphs of similar characters in the paragraph A) of the document WG2-N2644. The Editorial Group's response is that the similar characters identified in the document are not reflecting the complete picture adopting the same approach as in the ISO 10646. The document WG2-N2644 is using a **single** glyph for each character, whereas the ISO 10646 document allows multiple glyphs presentation for a single character.

When two ideographs have almost same structures and abstract shapes, and when it is unclear, under usual unification rules, whether those two should be unified or should not be unified, The Editorial Group had adopted the following 'dictionary_based_unification_rule' should apply:

'If two ideographs appear one standard dictionary, and the dictionary gives same definition and pronunciation to them, then those two ideographs should be unified'.

Discussion Item 2.

Action: China to check and report

The Editorial Group studied the different glyphs of similar characters in the paragraph B) of the document WG2-N2644. The Editorial Group wanted to bring to the WG2's attention that characters were mapped not by the glyphs.




The Editorial Group had checked that the characters stated in the Item 3 and 5, and agreed that the characters identified should have been unified. It was confirmed that the sources of the characters were of the same glyph, therefore they could be unified. Considering that any post-implementation changes might have impacts causing unforeseeable consequences, the Editorial Group proposed to keep the characters as they are now with no immediate change.

As for the Item 4, the proposed solution is acceptable by the IRG.

Discussion Item 3.

Action: the Chief Editor, TCA for item c)

The Editorial Group reviewed the 4 characters (285ED, 29FCE, 21BA7, 21E45) presented in the document N973. The Editorial Group arrived at the following conclusion.

- a) The glyph of the character 285ED in N973 is . China shall propose new glyph.
- b) The 'ISO 10646-2:2001' glyph  used for the character 29FCE in N973 is correct.
- c) The Group requested TCA to review on the glyphs of character (21BA7) then report the findings to the Chief Editor.
- d) China had gathered supporting information indicating that the glyph of the character 21E45 is . The Group accepted China's recommendation.

The Editorial Group further instructed the Chief Editor to take appropriate follow-up actions to see to the proper implementation of the decision.

Discussion Item 4.

Unanimous


The Editorial Group reviewed the characters information presented in the document N1006 submitted by Japan. The Editorial Group had decided to accept Japan’s recommendation that:

- a) the character 20F96 should have been unified to the character 5668.
- b) the character 22049 should ordinarily have been unified to the character 5DFD.
- c) the character 237EC should have been unified to the character 6B04.
- d) the character 24D01 should have been unified to the character 7589.
- e) the character 2592E should have been unified to the character 7A81 according to Kangxi.
- f) the character 28E93 should have been unified to the character 9686.
- g) the character 29516 should have been unified to the character 985E.

Discussion Item 5.

The Editorial Group reviewed the characters information presented in the document N972 submitted by China. The Editorial Group concluded that except the characters in the following table, the left glyph (Gly 128 01) are the preferred choice.

Unicode of Character	Preference by the IRG
2023D	No change
2042D	The glyph on right (Gly 128 03)
20BF6	Both glyphs are incorrect
20E03	No Change
21BA7	The font will be modified according to the proposed font by TCA
20FC6	No Change
21E45	The glyph on right (Gly 128 03)
25418	TCA explained that the origin should have the upper-most horizontal stroke is longer. Since character variants preferred should look more similar to the one in G_KX. The middle horizontal stroke of glyph on right should be shorten. The glyph on right (Gly 128 03)
2558D	The glyph on right (Gly 128 03)
25AED	The glyph on right (Gly 128 03)
25A4F	The glyph on right (Gly 128 03)
25A87	The glyph on right (Gly 128 03)
25CDE	The glyph on right (Gly 128 03)
2617E	The glyph on right (Gly 128 03)
261F6	The glyph on right (Gly 128 03)

2520D	The glyph on right (Gly 128 03)
25E12	The glyph on right (Gly 128 03)
2644A	The glyph on right (Gly 128 03)
2645A	The glyph on right (Gly 128 03)
26475	The glyph on right (Gly 128 03)
26486	The glyph on right (Gly 128 03)
26FC9	The glyph on right (Gly 128 03)
270E1	The glyph on right (Gly 128 03)
26597	The glyph on right (Gly 128 03)
2779D	The glyph on right (Gly 128 03)
2A6B2	The glyph on right (Gly 128 03)
28446	The glyph on right (Gly 128 03)
28597	The glyph on right (Gly 128 03)
285ED	Neither glyph of Gly128 01 nor Gly128 03. The glyph needs to be same as agreed in  (KX 1266.290)
28F32	The glyph on right (Gly 128 03)

Discussion Item 6.

Action: HKSAR

The Editorial Group reviewed the ‘Possible Errors in CJK, CJK_A and CJK_B information’ presented in the document N1008R submitted by China.

The Editorial Group reviewed the attachment ‘Duplication in CJK’. The Editorial Group considered that except the character pairs listed in Table A) and Table B) below, **the characters should have been unified, but decided not to fix** them at the moment.

Table A. The Editorial Group considered that the following characters **cannot be unified**.

UCS Code point	UCS Code point	Summary of discussion / Reason
07808	25449	They are considered non-cognate, since Kangxi, HYD and CNS separate them
08FFA	284BF	
20570	23353	SEPARATE per its TOP components
20566	2322F	
20560	0665F	
2055D	2319A	
06C77	23C8B	
25864	25871	For character 25864, G shape and T shape should have not been unified. Instead, the T shape should have been unified with G source shape of character 25871.
05B0E	05B14	Source code separation
04EDE	04EED	Source code separation
03B35	080F6	Different meanings in Kangxi

03B3B	04420	Different meanings in Kangxi
06727	268AB	Different meanings in Kangxi
23377	0813C	Different meanings in Kangxi

Table B. The Editorial Group decided that additional information are required / has follow-up actions.

UCS Code point	UCS Code point	Action
232AB	232AD	HKSAR explained that the two characters were different therefore cannot be unified. Japan opined that the two characters can ordinarily be unified therefore requested HKSAR to provide background information of the characters 232AB and 232AD for considerations.

The Editorial Group reviewed the attachments '*Maybe unified in CJK*'. The Editorial Group decided that the following character pairs listed in Table C) below should have been unified. The rest of the character pairs other than those listed in Table C) will stay as separate.

Table C. The following characters **SHOULD HAVE BEEN UNIFIED**.

UCS Code point	UCS Code point
259B1	25999
20163	21567
2462C	24631
24742	072C7
6267	22A7E
2866C	09094
204D3	28BA0

The Editorial Group started reviewing the attachment '*Unified errors in CJK*'. Due to limited time available, the Editorial Group instructed members to study the attachment and submit their comments to the Chief Editor before February 1, 2004.

IRG Editorial Group
19.11.2003

Attachment – Records of detailed discussion.

Review results on N1006

- 00005668 v.s. 00020F96 They should have been unified.
- 00005DFD v.s. 00022049 They should ordinarily have been unified.
- 00006B04 v.s. 000237EC They should have been unified.
- 00007589 v.s. 00024D01 They should have been unified.
- 00007A81 v.s. 0002592E They should have been unified, especially since
Kangxi says they are same.
- 00009686 v.s. 00028E93 They should have been unified.
- 0000985E v.s. 00029516 Based on the glyphs, they should have been unified.

Review results on N1008R

- 00021BC3 v.s. 0000352B Glyph shapes suggests they should have been unified,
but they may have been deunified for some other reason.
They should be kept as is, since they have their own
published mappings.
- 00007808 v.s. 00025449 They are considered non-cognate, since Kanxi, HYD and
CNS separate them.
- 00008FFA v.s. 000284BF Their shapes look unifiable. However, since HYD gives
different meanings and pronunciation to them,
non-cognate rule apply. They should not be unified.
- 00027729 v.s. 00004666 Unifiable shapes. Same meaning and sounds in Kangxi.
They should have been unified.
- 00020570 v.s. 00023353 SEPARATE per its TOP components.
- 000232AB v.s. 000232AD Glyph shapes look unifiable. HK believes they have
different meanings (non-cognate.) IRG wants further
input from HK.
- 00020566 v.s. 0002322F SEPARATE as 00020570 v.s. 00023353
- 00020560 v.s. 0000665F SEPARATE as 00020570 v.s. 00023353
- 0002055D v.s. 0002319A SEPARATE as 00020570 v.s. 00023353

00023F41 v.s. 00023F9E	should have been unified.
00023F4D v.s. 00023F84	should have been unified.
00006F78 v.s. 00023F7D	should have been unified
00006C77 v.s. 00023C8B	should have been unified
0002633D v.s. 00027B80	should have been unified.
00026288 v.s. 0002668B	Meat v.s. Moon issue. Postpone.
00007468 v.s. 00024A01	should have been unified
00028521 v.s. 00028535	should have been unified
00008A04 v.s. 00027960	difference in structure? Just the length?
00021A0D v.s. 00024B3D	should have been unified
-	
000216AE v.s. 000216B0	should have been unified
000038AA v.s. 000223B0	should have been unified
00025864 v.s. 00025871	For 00025864, G shape and T shape should have not been unified. Instead, the T shape should have been unified with G source shape of 00025871.
000085BB v.s. 000270E1	should have been unified.
00027144 v.s. 00008641	should have been unified.
00025C83 v.s. 00007C06	should have been unified.
00005B0E v.s. 00005B14	Source code separation. (Note that both are in BMP)
00021018 v.s. 0002103C	should have been unified.
0002243D v.s. 000263DA	TCA to check
00004EDE v.s. 00004EED	Source code separation.
00003BA3 v.s. 000069E9	should have been unified.
00003588 v.s. 0000439B	should have been unified.
0000363D v.s. 000039B3	should have been unified.

00023761 v.s. 000237BC should have been unified.

00003DB7 v.s. 0002420E should have been unified.

00003B35 v.s. 000080F6 non-cognate, especially since Kangxi gives different meanings.

00008192 v.s. 000267FE should have been unified.

00003B3B v.s. 00004420 non-cognate, especially since Kangxi gives different pronunciation and meanings.

00006727 v.s. 000268AB non-cognate, especially since Kangxi gives different pronunciation.

00023377 v.s. 0000813C non-cognate, especially since Kangxi gives different pronunciation.

000043DE v.s. 0002669C should have been unified.

00005E50 v.s. 00003B3A Glyph shapes look appropriate for unification. Although these two are classified under different radicals, their pronunciation and meanings are same, and they are **not** considered to be non-cognate. So, IRG Editors considers these two should have been unified.

-

000219AD v.s. 000219B4 Not suitable for unify in their glyph shapes.

00028B27 v.s. 00028B56 Not suitable for unify in their glyph shapes.

00020505 v.s. 00021C24 Glyph shapes appear to be unifiable, but Kangxi shows different meanings, so they are considered non-cognate.

00003563 v.s. 00020BA6 Non-cognate since HYD says they are different.

00021D00 v.s. 00021D44 Not suitable for unify, since their stroke counts are different.

000284AF v.s. 000284BD China to check HYD

00021E13 v.s. 00021E9B China to check HYD

00021A40 v.s. 00021A5A China to check HYD

000259B1 v.s. 00025999 Should have been unified.

00020824 v.s. 00020825 China to check HYD

000250E2 v.s. 000268F8 China to check HYD

000083FB v.s. 00026C4D Glyph shapes look unifiable, but Kangxi says they
are different. Non-cognate rule apply.

00020163 v.s. 00021567 should have been unified.

0002462C v.s. 00024631 should have been unified.

00024742 v.s. 000072C7 should have been unified

00004F35 v.s. 000201FB Non-cognate, since Korean dictionary says they have
totally different meanings.

00006267 v.s. 00022A7E should have been unified.

000206C3 v.s. 00022A6C Comparing glyphs shapes, unifiability of these two
is unambiguous. Some of editors feel they are
different enough, and some other feel they can be
unifiable. China to check HYD.

0002866C v.s. 00009094

00027DA0 v.s. 00008CE3 Although their shapes look close enough to be unified,
Kangi says they are different. Non-cognate rule apply.

000290AB v.s. 000290AD Lower omponents of these two are too different to unify.

00006747 v.s. 000233D3 Non-cognate characters, since in Japan they have totally
different meanings.

000204D3 v.s. 00028BA0 They should have been unified, since IRG Editors consider that
the difference is only stylistic (i.e., the difference in the length of the last stroke) and is *not*
the difference in overall structure.

00022189 v.s. 0002634D Not to be unified.

00023385 v.s. 00026772 Non-cognate, since Kangxi gives different meanings.

0000670A v.s. 000043D3 Non-cognate, since Kangxi gives different meanings.

0002339B v.s. 00026808 Non-cognate, since Kangxi gives different meanings.

0002339F v.s. 00026805 Glyph shapes look same, but they are considered
non-cognate per TCA's statements on the origin of
the characters.

00006726 v.s. 00004443 Non-cognate, since Kangxi gives different meanings.

00003B36 v.s. 000266E9 Glyph shapes look same, but one is from meat radical and another from moon radical. They have different meanings. So, they are considered non-cognate.

00023366 v.s. 0002667F Glyph shapes look same, but very likely to be one is from meat radical and another from moon radical.
TCA to check

-