

IRG N1838
















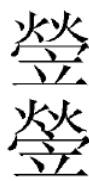
Title: CJK B error report to WG2







Source: IRG editorial group

Meeting: IRG#37

Date: 2011-11-10





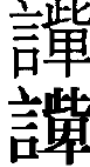
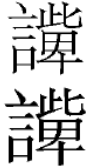
IRG found several errors on CJK Unified Ideographs Extension B while [doing](#) horizontal extension work. Some of them are very serious and need to be known widely by UCS users. IRG requests to include these errata in nearest issue of [the](#) errata report of UCS.

| Unicode   | UCS 2003 Glyph   | IRG N1707 Glyph   | SuperCJK Glyph  |
|---|--|---|---|
| 2382C   | <br>UCS-2382C   | 2382C<br>木 75.18<br> TF-6951<br> UCS-2382C  |       |
| TF-6951 was mistakenly unified to U+2382C at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B. TF-6951 of ISO/IEC10646 keeps this shape for the consistency to the CNS standard although these two glyph shapes cannot be unified.   |  |   |   |
| 23EE4   | <br>UCS-23EE4 | 23EE4<br>水 85.11<br> GKX-0648.09<br> T7-243F<br> UCS-23EE4 |   |
| T7-243F was mistakenly unified to U+23EE4 at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B. T7-243F of ISO/IEC10646 keeps this shape for the consistency to the CNS standard although these two glyph shapes cannot be unified.<br>[Note] GKX-0648.09 can be unified to U+23EE4 although it has slight difference from KangXi dictionary shape. |  |   |   |
| 24369   | <br>UCS-24369 | 24369<br>火 86.12<br> TF-5024<br> UCS-24369  |   |
| Because of mis-unification in Super CJK, UCS glyph was designed differently (one dot more) from TF-5024. To keep consistency multicolumn code chart keeps original UCS shape.   |  |   |   |






| Unicode | UCS 2003 Glyph   | IRG N1707 Glyph   | SuperCJK Glyph   |
|---------|--|---|--|
| 27555   | <br>UCS-27555 | 27555<br>虫 142.17<br><br>GKX-1103.29<br><br>T5-7649<br><br>UCS-27555 | <br> |

Because of mis-unification in Super CJK, UCS glyph was designed differently (one dot more) from Kangxi dictionary shape. To keep consistency multicolumn code chart keeps original UCS shape.






[Note] UCS shape is wrong because GKX-1103.29 is placed at 17 strokes place in Kangxi dictionary.

|       |  |  |  |
|-------|--|--|--|
| 27B1F | <br>UCS-27B1F | 27B1F<br>言 149.12<br><br>GHZ-64018.09<br><br>T7-5035<br><br>UCS-27B1F | <br> |
|-------|--|--|--|






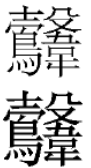
GHZ-65018.09 was mistakenly unified to U+27B1F at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B. GHZ-65018.09 of ISO/IEC10646 keeps this shape for the consistency of source mapping although these two glyph shapes cannot be unified.

|       |  |  |  |
|-------|--|--|--|
| 27D41 | <br>UCS-27D41 | 27D41<br>貝 154.4<br><br>TF-385F<br><br>UCS-27D41 | <br> |
|-------|--|--|--|

Because of mis-unification in Super CJK, UCS glyph was designed differently from TF-385F. To keep consistency multicolumn code chart keeps original UCS shape.

|       |  |   |  |
|-------|--|---|--|
| 28B75 | <br>UCS-28B75 | 28B75<br>金 167.13<br><br>TF-686D<br><br>UCS-28B75 | <br> |
|-------|--|---|--|

Shape of TF-686D in CNS standard has been changed after the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B was published, so TF-686D needs to keep this shape although it cannot be unified to U+28B75.






|       |  |  |  |
|-------|--|--|--|
| 293FB | <br>UCS-293FB | 293FB<br>草 178.20<br><br>GHZ-74512.13<br><br>T5-7C22<br><br>UCS-293FB | <br> |
|-------|--|--|--|

Shape of T5-7C22 in CNS standard has been changed after the first edition of CJK UNIFIED

| Unicode | UCS 2003 Glyph | IRG N1707 Glyph | SuperCJK Glyph |
|---------|----------------|-----------------|----------------|
|---------|----------------|-----------------|----------------|





IDEOGRAPHS EXTENSION B was published, so T5-7C22 needs to keep this shape although it cannot be unified to U+293FB.

[Comment] IRG didn't agree to discuss about modifying unification rule. So this is not an "exception" of the rule.





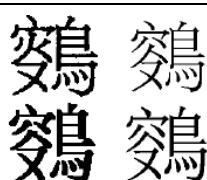
|       |  |  |   |
|-------|--|--|---|
| 29450 | <br>UCS-29450 | 29450<br>頁 181.3<br><br>GKX-1399.08<br><br>T5-3D4A<br><br>UCS-29450 |  |
|-------|--|--|---|

GKX-1399.08 has different shape from Kangxi Dictionary. Both GKX-1399.08 and T5-3D4A can be unified to U+29450.

[Comment] This note is not necessary because this is a case of issue 2 of 4.3 in IRG N1745.

|       |  |   |   |
|-------|--|---|---|
| 29C52 | <br>UCS-29C52 | 29C52<br>鬻 193.12<br><br>T7-5666<br><br>UCS-29C52 |  |
|-------|--|---|---|

Shape of T7-5666 in CNS standard has been changed after the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B was published, so T7-5666 needs to keep this shape although it cannot be unified to U+29C52.

|       |  |  |   |
|-------|--|--|---|
| 2A0B8 | <br>UCS-2A0B8 | 2A0B8<br>鷓 196.9<br><br>GKX-1494.15<br><br>T7-523A<br><br>UCS-2A0B8 |  |
|-------|--|--|---|

Shape of T7-523A in CNS standard has been changed after the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B was published, so T7-523A needs to keep this shape although it cannot be unified to U+2A0B8.






[Comment] Glyph difference of G is a popular pattern of unification but T is not, however TCA wants to keep this shape. So IRG concluded to make note for this situation. If TCA needs further study, IRG has better to discuss at the next meeting.

One more note should be added (cf. 2.2(h))

|       |  |  |   |
|-------|--|--|---|
| 299FB | <br>UCS-299FB | 299FB<br>马 187.8<br><br>GCH<br><br>UCS-299FB |  |
|-------|--|--|---|





Shape of GCH has been changed after the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B was published, so G-glyph for U+299FB needs to keep this shape although it

cannot be unified to the original UCS shape.

|         |  |                               |  |  |             |  |                   |
|---------|--|-------------------------------|--|--|-------------|--|-------------------|
| u+235f1 | <br>235F1 | 235F1<br>木 75.10<br>梨 UCS2003 | <br>UCS2003 | <br>V0-354D | wrong glyph | J: This is postponed at IRG #36 however they cannot be unified.<br>G:  and  not unifiable, need V glyph to change back..<br>K: V glyph wrong, an extra 人 is added. | IRG N1775 table 3 |
|---------|--|-------------------------------|--|--|-------------|--|-------------------|




No change. UCS2003 is incorrect. It is a unification error

V0-354d was mistakenly unified to u+235f1 at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B. IRG concluded that V0-354d of ISO/IEC 10646 keeps this shape for the consistency to the TCVN5773 although these glyphs are not able to be unified.

|         |  |                                |  |  |  |             |   |                             |
|---------|--|--------------------------------|--|--|--|-------------|---|-----------------------------|
| u+28599 | <br>28599 | 28599<br>辶 162.11<br>遯 UCS2003 | <br>UCS2003 | <br>T5-519D | <br>V4-5565 | unification | J: This is concluded at IRG #36 that V to change back to UCS2003 shape.<br>H: The V-glyph has been modified, but the new glyph has one additional stroke.<br>G: An extra stroke been added compares to UCS font, unifiable?<br>K: non-unifiable shape | IRG N1792/IRG N1775 table 1 |
|---------|--|--------------------------------|--|--|--|-------------|---|-----------------------------|




No change. Make a note to corrigenda as a unification error.

V4-5565 was mistakenly unified to u+28599 at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B. IRG concluded that V4-5644 of ISO/IEC 10646 keeps this mapping for the consistency with the V4 source although these glyphs are not able to be unified.

|         |   |                               |  |   |  |                   |                      |
|---------|---|-------------------------------|--|---|--|-------------------|----------------------|
| U+20885 | T | 20885<br>力 19.10<br>勅 UCS2003 | <br>UCS2003 | <br>GHOX-0149.28 | <br>T5-3669 | glyph not revised | IRG N1775 table4-2.1 |
|---------|---|-------------------------------|--|---|--|-------------------|----------------------|

T-glyph can be unified to U+2088D. But we would not change the map. It is a mapping error.



T5-3669 was mistakenly unified to u+20885 at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B. IRG concluded that IRG concluded that T5-3669 should be mapped to u+2088d by means of unification rule and this mapping is err. However, IRG concluded that T5-3669 of ISO/IEC 10646 keeps this mapping for the consistency with CNS standard.

|         |   |                               |  |   |  |  |                         |
|---------|---|-------------------------------|--|---|--|--|-------------------------|
| U+21F12 | T | 21F12<br>山 46.12<br>嶽 UCS2003 | <br>UCS2003 | <br>GHOX-0319.30 | <br>T5-4C6E | J: T source should be removed for a new character as explained in IRG N1793.<br>K: T glyph is the same as 21F2C□ | IRG N1775 table4-2.2(c) |
|---------|---|-------------------------------|--|---|--|--|-------------------------|



Move T-source from U+21F12 to U+21F2C. Record U-glyph of U+21F2C is wrong. It is a mapping error.

T5-4C6E was mistakenly unified to u+20885 at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B.



IRG concluded to ask WG2 moving T5-4C6E source to map to u+21f2c.

|  |   |  |  |  |                            |  |
|--|---|--|--|--|----------------------------|--|
| U+24A8A  | V | 24A8A<br>五 96.13<br><br>UCS2003 | <br>V2-7C66 | J: glyph not revised<br>H: The V-glyph has 12 strokes, not 13 as in the UCS2003 glyph. | IRG N1775<br>table4-2.2(f) |  |
| No change. UCS is wrong. It is a unification error |   |  |  |  |                            |  |




Shape of u+2a48a was wrong in ISO/IEC10646:2003. Although correct shape of V2-7C66 cannot be unified to the current shape in ISO/IEC10646:2003, IRG concluded to keep this mapping because of the consistency.

|  |   |   |  |   |                            |  |
|--|---|---|--|---|----------------------------|--|
| U+24F15  | V | 24F15<br>フ 104.18<br><br>UCS2003 | <br>V2-7D5A | J: glyph not revised<br>G: A stroke is missing in V glyph. Already mentioned in IRGN1775_A4 (J comments).<br>K: wrong glyph | IRG N1775<br>table4-2.2(f) |  |
| No change. UCS is wrong. It is a unification error |   |   |  |   |                            |  |





Shape of u+24f15 was wrong in ISO/IEC10646:2003. Although correct shape of V2-7D5A cannot be unified to the current shape in ISO/IEC10646:2003, IRG concluded to keep this mapping because of the consistency.

|   |   |   |  |                   |                         |  |
|---|---|---|--|-------------------|-------------------------|--|
| U+25089   | V | 25089<br>血 108.10<br><br>UCS2003 | <br>V2-7D68 | glyph not revised | IRG N1775 table4-2.2(f) |  |
| No change. UCS is wrong. It is a unification error. |   |   |  |                   |                         |  |

Shape of u+25089 was wrong in ISO/IEC10646:2003. Although correct shape of V2-7D68 cannot be unified to the current shape in ISO/IEC10646:2003, IRG concluded to keep this mapping because of the consistency.





|  |   |  |  |  |                   |                    |  |
|--|---|--|--|--|-------------------|--------------------|--|
| U+2A6C0  | G | 2A6C0<br>龜 213.8<br><br>UCS2003 | <br>GKX-1538.20 | <br>T5-7B5E | glyph not revised | IRG N1775 table4-3 |  |
| No change and add a note to corrigenda as Unification error. |   |  |  |  |                   |                    |  |

GKX-1538.20 was mistakenly unified to u+2a6c0 at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B. V4-5644 of ISO/IEC 10646 keeps this shape for the consistency to the GKX source although these glyphs are not able to be unified.



|       |  |  |  |  |             |  |  |
|-------|--|--|--|--|-------------|--|--|
| 22936 | <br>UCS2003 | <br>UCS2003 | <br>GKX-0408.28 | <br>T5-6777 | Unification | The T-glyph is different from the KX Dictionary glyph. |  |
|-------|--|--|--|--|-------------|--|--|

No change. Add a note to corrigenda as a unification error



T5-6777 was mistakenly unified to u+22936 at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B. T5-6777 of ISO/IEC 10646 keeps this shape for the consistency to the CNS standard although these glyphs are not able to be unified.

|  |  |  |  |  |             |   |  |
|--|--|--|--|--|-------------|---|--|
| 28321  | <br>UCS2003 | <br>UCS2003 | <br>GKX-1244.18 | <br>T6-632A | Unification | The T-glyph is different from the <i>KX Dictionary</i> glyph. |  |
| No change and add a note to corrigenda as a unification error. |  |  |  |  |             |   |  |



T6-632A was mistakenly unified to u+22936 at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B. T6-632A of ISO/IEC 10646 keeps this shape for the consistency to the CNS standard although these glyphs are not able to be unified.

|  |  |  |             |   |   |
|--|--|--|-------------|---|---|
| 22BA3  | <br>UCS2003 | <br>GKX-0440.17 | Wrong glyph | Difference in the top part of the glyph | G |
| No change and add a note to corrigenda as a unification error. |  |  |             |   |   |

GKX-0440.17 was mistakenly unified to u+22BA3 at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B. GKX-0440.17 of ISO/IEC 10646 keeps this shape for the consistency to the GKX source although these glyphs are not able to be unified.

|  |  |  |             |  |   |
|--|--|--|-------------|--|---|
| 23023  | <br>UCS2003 | <br>T5-6C34 | Wrong glyph | Difference in the right component of the glyph | T |
| No change and add a note to corrigenda as a unification error. |  |  |             |  |   |

T5-6C34 was mistakenly unified to u+23023 at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B. T5-6C34 of ISO/IEC 10646 keeps this shape for the consistency to the CNS standard although these glyphs are not able to be unified.

|  |  |  |             |                       |   |
|--|--|--|-------------|-----------------------|---|
| 24229  | <br>UCS2003 | <br>GKX-0672.02 | Wrong glyph | The bottom difference | G |
| No change and add a note to corrigenda as a unification error. |  |  |             |                       |   |

GKX-0672.02 was mistakenly unified to u+23023 at the first edition of CJK UNIFIED IDEOGRAPHS EXTENSION B. GKX-0672.02 of ISO/IEC 10646 keeps this shape for the consistency to the GKX source although these glyphs are not able to be unified.

(End of Document)