**Doc Type:** Working Group Document

Comments on nine error glyphs, three wrong glyph designs & one unification

Title:

on CJK Unified Ideographs Ext B

Source: Eiso Chan (陈永聪), Victor Wong

Author: Eiso Chan (陈永聪)

**Status:** Individual Contribution

Action: For consideration by JTC1/SC2/WG2/IRG and UTC

Date: 2017-03-12

#### 1. Possible error V-source glyph for U+21846

The IDS of the current UCS2003 glyph for U+21846 is  $\square$  女眼, but the IDS of the current V-source glyph for U+21846 is  $\square$  女目? . (pls see Fig. 1.1)  $\square$  目? shouldn't be unifiable with 眼.

Fig. 1.1 U+21846 on The Unicode Standard, Version 9.0

We found this character in the offical website of VNPF, Quốc Ngữ is mìn. (pls see Fig. 1.2)

Fig. 1.2 U + 21846 on VNPF

| Quốc Ngữ | Hán-Nôm | Codepoint      | Context         | Ref. | English         |
|----------|---------|----------------|-----------------|------|-----------------|
| mìn      | 娜       | <u>U+21846</u> | mẹ mìn, mìn mịn | vhn  | child kidnapper |

#### Fig. 1.3 U+21846 on Kangxi-Plus Unicode Han Dictionary

张玉书, 陈廷敬, 王宏源 等 (Zhāng Yùshū, Chén Tíngjìng, Wáng Hóngyuán, et al): 《康熙字典(增订版)》(*Kangxi-Plus Unicode Han Dictionary*), 北京: 社会科学文献出版社 [Běijīng: Social Sciences Academic Press (CHINA)], 2015, ISBN 978-7-5097-7093-1, p. 302

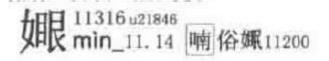


Fig. 1.4 U + 2A993 on Kangxi-Plus Unicode Han Dictionary

张玉书, 陈廷敬, 王宏源 等 (Zhāng Yùshū, Chén Tíngjìng, Wáng Hóngyuán, et al): 《康熙字典(增订版)》(*Kangxi-Plus Unicode Han Dictionary*), 北京: 社会科学文献出版社 [Běijīng: Social Sciences Academic Press (CHINA)], 2015, ISBN 978-7-5097-7093-1, p. 300

According to Fig. 1.3, U+21846 is the variant form for U+2A993. According to Fig. 1.4, the semantic component (形符) of U+2A993 is  $\pm$ , and the other part is the phono component (声符). Quốc Ngữ of 眠 is miên which is similar with mìn. (pls see Fig. 1.5) So the best form for mìn in Vietnamese word mẹ mìn & mìn mịn is U+2A993. The right hand side of UCS2003 & V-source glyphs are all the variant form for 眠, but they couldn't be unified with each other.

Fig. 1.5 U + 7720 on VNPF

| Quốc Ngữ | Hán-Nôm | Codepoint     | Context   | Ref. | English                    |
|----------|---------|---------------|-----------|------|----------------------------|
| miên     | 眠       | <u>U+7720</u> | thôi miên | vhn  | to hypnotize, to mesmerize |

We suggest three ways to solve this problem:

A1) Encode V0-3521 as a new character, and the kRSUnicode value for the new character is 38.11.

A2) Add a new U-source glyph into U+21846. The evidence of the new U-source ideograph is

Fig. 1.3.

Table 1.1 Information for a New Proposed U-source Ideograph

| Glyph | IDS                                 | Radical | Radical<br>Form | Strokes | First<br>Stroke | Simplified<br>Form | Variant<br>Form | Variant<br>Code Point | U-source<br>Number     |
|-------|-------------------------------------|---------|-----------------|---------|-----------------|--------------------|-----------------|-----------------------|------------------------|
| 艉     | □女眼<br>U+2FF0,<br>U+5973,<br>U+773C | 38      | 女               | 11      | 2               | N/A                |                 | U+2A993               | UTC-xxxxx OR UCI-xxxxx |

B) Keep the present situation of U+21846, and add two glyphs into SVS like this:

Table 1.2 Proposed SVSes for U+21846

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+21846        | FE00 (VS1)         | 艉     | UCS2003     |
| U+21846        | FE01 (VS2)         | 覷     | V0-3521     |

C) Keep the present situation of U + 2017B, and add two glyphs into IVD like this:

Table 1.3 Proposed IVSes for U+21846

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+21846        | E0100 (VS17)       | 艉     | UCS2003     |
| U+21846        | E0101 (VS18)       | 娜     | V0-3521     |

## 2. Wrong V-source glyph design for U+21BC0

Fig. 2.1 U+21BC0 on The Unicode Standard, Version 9.0



J-source, G-source, H-source & K-source had pointed out this mistake to V-source in IRG #37, and V-source answered that the V-glyph would be updated. (pls see <a href="IRGN1784VietnamResponse">IRGN1784VietnamResponse</a> or Fig. 2.2)

Fig. 2.2 IRGN1784VietnamResponse, p. 2

| u+21bc0 | <b>性</b> 图 | 21BC0<br>1 4229 | 鬱\<br>UCS2003 | 特林<br>世<br>197<br>V2-7456 | wrong glyph | J: V2-7456 is modified as in IRG N1792<br>but it is still wrong.<br>G: A horizontal stroke missing, should be<br>corrected.<br>H: The V-glyph has been modified, but<br>one stroke is missing in the new glyph.<br>K: Four dots are missing in the V-glyph. | V-font will be updated | 檀沙 |
|---------|------------|-----------------|---------------|---------------------------|-------------|---|------------------------|----|
|---------|------------|-----------------|---------------|---------------------------|-------------|---|------------------------|----|

We suggest that V-source update V-glyph as soon as possible.

#### 3. Error G-source glyph for U + 2278B

The IDS of the current UCS2003 glyph for U+2278B is  $\Box \uparrow \Box f \Box f$ , but the IDS of the current G-source glyph for U+2278B is  $\Box \uparrow \Box f \Box f$ . (pls see Fig. 3.1)  $\uparrow$  shouldn't be unifiable with  $\uparrow$ .

Fig. 3.1 U + 2278B on The Unicode Standard, Version 9.0

We showed the evidence in G-sources: GHZ-10207.12 which is in the 《汉语大字典》. (pls see Fig. 3.2)

Fig. 3.2 UCS2003 glyph for U+2278B on Hànyǔ Dàzìdiǎn V1, p. 207

**伝** tǔn《字彙補・心部》:"億,透混切。見《皇極圖 韻》。"

The glyph in 《汉语大字典》 V2 is the same. (pls see Fig. 3.3)

#### Fig. 3.3 UCS2003 glyph for U+2278B on Hànyǔ Dàzìdiǎn V2

汉语大字典编辑委员会(Hànyǔ Dàzìdiǎn Biānjí Wěiyuánhuì):《汉语大字典(第二版)》(*Hànyǔ Dàzìdiǎn V2*), 武汉: 湖北长江出版集团崇文书局 (Wǔhàn: Chongwen Publishing House of Hubei Changjiang Publishing Group) & 成都: 四川出版集团四川辞书出版社 (Chéngdū: Sichuan Reference Press of Sichuan Publishing Group), 2010, ISBN 978-7-5403-1744-7, p. 246

(香 tǔn 《字彙補·心部》: "億,透混切。見《皇極圖韻》。"

We have found more evidences about this glyph.

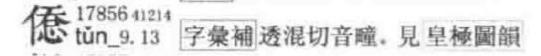
Figs. 3.4 through 3.6 UCS2003 glyph for U+2278B on different versions of KangXi Dictionary 内府本《康熙字典》, p. 3725



同文書局版《康熙字典》, p. 1555



张玉书, 陈廷敬, 王宏源 等 (Zhāng Yùshū, Chén Tíngjìng, Wáng Hóngyuán, et al): 《康熙字典(增订版)》(*Kangxi-Plus Unicode Han Dictionary*), 北京: 社会科学文献出版社 [Běijīng: Social Sciences Academic Press (CHINA)], 2015, ISBN 978-7-5097-7093-1, p. 479



We suggest four ways to solve this problem:

- A) Ask G-source to change the GHZ-10207.12 glyph for U + 2278B.
- B1) Withdraw GHZ-10207.12 source glyph for U+2278B.
- B2) Add a new U-source glyph into U+2278B. The evidence of the new U-source ideograph is

Figs. 3.2 through 3.6.

Table 3.1 Information for a New Proposed U-source Ideograph

| Glyph | IDS   | Radical | Radical<br>Form | Strokes | First<br>Stroke | Simplified<br>Form | Variant<br>Form | Variant<br>Code Point | U-source<br>Number     |
|-------|---|---------|-----------------|---------|-----------------|--------------------|-----------------|-----------------------|------------------------|
| 恁     | U+2FF0,<br>U+4EBB,<br>U+2FF1<br>U+541E,<br>U+5FC3 | 61      | 心               | 9       | 3               | N/A                | N/A             | N/A                   | UTC-xxxxx OR UCI-xxxxx |

C) Keep the present situation of U + 2278B, and add two glyphs into SVS like this:

Table 3.2 Proposed SVSes for U + 2278B

| Base Character | Variation Selector | Glyph | Description  |
|----------------|--------------------|-------|--------------|
| U+2278B        | FE00 (VS1)         | 億     | UCS2003      |
| U+2278B        | FE01 (VS2)         | 憶     | GHZ-10207.12 |

D) Keep the present situation of U + 2278B, and add two glyphs into IVD like this:

Table 3.3 Proposed IVSes for U + 2278B

| Base Character | Variation Selector | Glyph | Description  |
|----------------|--------------------|-------|--------------|
| U+2278B        | E0100 (VS17)       | 億     | UCS2003      |
| U+2278B        | E0101 (VS18)       | 憶     | GHZ-10207.12 |

## 4. Possible error T-source glyph for U + 22E3F

The IDS of the current UCS2003 glyph for U+22E3F is □ ‡ 篚, but the IDS of the current T-source glyph for U+22E3F is □ 木篚. (pls see Fig. 4.1) 木 shouldn't be unifiable with ‡.

Fig. 4.1 U+22E3F on The Unicode Standard, Version 9.0

The official website of the Master Ideographs Seeker for CNS11643 Chinese Standard Interchange Code (CNS11643 中文標準交換碼全字庫) shows this contradiction because of the disunification of two glyphs for U+22E3F. (pls see Fig. 4.2)

Fig. 4.2 U + 22E3F on The Master Ideographs Seeker

| 全字庫字型        |   | 發音        |      |  |  |  |  |  |
|--------------|---|-----------|------|--|--|--|--|--|
| 推推 權明體 指體 宋體 | □へ。<br>男聲 <mark>◆</mark> 於女 <mark>聲◆</mark> 於 |           |      |  |  |  |  |  |
| 拼音           |   | 漢語<br>fěi |      |  |  |  |  |  |
| 字義           | 筆畫 部首 倉頡                                      |           |      |  |  |  |  |  |
| 相關詞          | 19  | 手         | QHSY |  |  |  |  |  |
| 部 件          | 末, 体, 匚, 非, 部件版本(CNS11643_2_94)               |           |      |  |  |  |  |  |
| 筆順序          | , ヿ , ー , ー , ー , ー , ー , ー , ー , ー , ー       |           |      |  |  |  |  |  |
| 漢字國屬性        | 台灣(TF-6246)                                   |           |      |  |  |  |  |  |

We could find out that the current T-source glyph for U+237E2 (aka TF-653D) is the same as the current T-source glyph for U+22E3F (aka TF-6246).

Fig. 4.3 U+237E2 on The Unicode Standard, Version 9.0

We suggest four ways to solve this problem:

- A) Ask T-source to change the TF-6246 glyph for U + 22E3F.
- B1) Withdraw the TF-6246 glyph for U + 22E3F.

B2) Add a new U-source glyph into U+22E3F.

Table 4.1 Information for a New Proposed U-source Ideograph

| Glyph | IDS                                 | Radical | Radical<br>Form | Strokes | First<br>Stroke | Simplified<br>Form | Variant<br>Form | Variant<br>Code Point | U-source<br>Number |
|-------|-------------------------------------|---------|-----------------|---------|-----------------|--------------------|-----------------|-----------------------|--------------------|
| 攤     | □才篚<br>U+2FF0,<br>U+624C,<br>U+7BDA | 64      | 手               | 16      | 3               | N/A                | N/A             | N/A                   | UCI-xxxxx          |

- C1) Keep the present situation of U+22E3F, and add a second kRSUnicode value of 75.16 in the future.
- C2) Add two glyphs into SVS like this:

Table 4.2 Proposed SVSes for U + 22E3F

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+22E3F        | FE00 (VS1)         | 攤     | UCS2003     |
| U+22E3F        | FE01 (VS2)         | 權     | TF-6246     |

- D1) Keep the present situation of U+22E3F, and add a second kRSUnicode value of 75.16 in the future.
- D2) Add two glyphs into IVD like this:

Table 4.3 Proposed IVSes for U + 22E3F

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+22E3F        | E0100 (VS17)       | 攤     | UCS2003     |
| U+22E3F        | E0101 (VS18)       | 權     | TF-6246     |

## 5. Error V-source glyph for U+235F1

The IDS of the current UCS2003 glyph for U+235F1 is  $\square$  梨下, but the IDS of the current V-source glyph for U+235F1 is  $\square$  国利人木下. (pls see Fig. 5.1) 国利人木 shouldn't be unifiable with 梨.

Fig. 5.1 U+235F1 on The Unicode Standard, Version 9.0

235F1 木 75.10 **秋** UCS2003 V0-354D

We found this character in the offical website of VNPF, Quốc Ngữ is lè. (pls see Fig. 1.2)

Fig. 5.2 U + 235F1 on VNPF

| Quốc Ngữ | Hán-Nôm | Codepoint      | Context    | Ref. | English |
|----------|---------|----------------|------------|------|---------|
| lè       | 梨下      | <u>U+235f1</u> | thấp lè tè | vhn  |         |

Fig. 5.3 U + 68A8 on VNPF

| Quốc Ngữ | Hán-Nôm | Codepoint     | Context | Ref. | English                 |
|----------|---------|---------------|---------|------|-------------------------|
| lê       | 梨       | <u>U+68a8</u> | quả lê  | vhn  | pear; opera; cut, slash |

We suggest four ways to solve this problem:

A) Change the V-source glyph to □梨下 like the glyph on VNPF. (pls see <u>IRGN1784</u>

## VietnamResponse or Fig. 5.4)

Fig. 5.4 IRGN1784 VietnamResponse

| Code    | Glyph<br>in<br>UCS<br>2003 |                  | oh in IR      | GN1784        | Comment<br>type | Comment  | V-Answer               | V-Font |
|---------|----------------------------|------------------|---------------|---------------|-----------------|--|------------------------|--------|
| u+235f1 | 梨下<br>235F1                | 235F1<br>* 75.10 | 梨下<br>UCS200S | 梨T<br>V0-354D | wrong glyph     | J: This is postponed at IRG #36 however they cannot be unified.    J   J | V-font will be updated | 楘      |

- B1) Encode V0-354D as a new character, and the kRSUnicode value for the new character is 75.12.
- B2) Add a new U-source glyph into U+235F1.

Table 5.1 Information for a New Proposed U-source Ideograph

| Gly | ph | IDS                                 | Radical | Radical<br>Form | Strokes | First<br>Stroke | Simplified<br>Form | Variant<br>Form | Variant<br>Code Point | U-source<br>Number |
|-----|----|-------------------------------------|---------|-----------------|---------|-----------------|--------------------|-----------------|-----------------------|--------------------|
| 利米  | F  | □梨下<br>U+2FF0,<br>U+68A8,<br>U+4E0B | 75      | 木               | 10      | 3               | N/A                | N/A             | N/A                   | UCI-xxxxx          |

- C1) Keep the present situation of U+235F1, and add a second kRSUnicode value of 75.12 in the future.
- C2) Add two glyphs into SVS like this:

Table 5.2 Proposed SVSes for U + 235F1

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+235F1        | FE00 (VS1)         | 梨     | UCS2003     |
| U+235F1        | FE01 (VS2)         | 梨     | V0-354D     |

D1) Keep the present situation of U+235F1, and add a second kRSUnicode value of 75.12 in the future.

# D2) Add two glyphs into IVD like this:

Table 5.3 Proposed IVSes for U + 235F1

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+235F1        | E0100 (VS17)       | 絮     | UCS2003     |
| U+235F1        | E0101 (VS18)       | 楘     | V0-354D     |

## 6. Error UCS2003 glyph for U+2386E

The IDS of the current UCS2003 glyph for U+2386E is  $\square$  ‡ 欎, but the IDS of the current T-source glyph for U+2386E is  $\square$  木 豑. (pls see Fig. 6.1) 木 shouldn't be unifiable with  $\ddagger$ .

Fig. 6.1 U+2386E on The Unicode Standard, Version 9.0

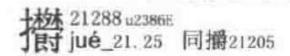
Fig. 6.2 U + 2386E on The Master Ideographs Seeker

| 全字庫字型        |  | 發音   |             |  |  |  |
|--------------|--|--|-------------|--|--|--|
| <b>樹 楷 楷</b> |  | 4□世 <sup>*</sup><br>男 <b>聲◆</b> 修女 <b>聲◆</b> 修 |             |  |  |  |
| 拼音           |  | 漢語<br>jué                                      |             |  |  |  |
| 字義           | 筆畫   | 部首   | 倉頡          |  |  |  |
| 相關詞          | 25   | 木  | DDDI        |  |  |  |
| 部件           | 木,木,夕,木,宀  | 艮,寸,部件版本(CNS                                   | 11643_2_94) |  |  |  |
| 筆順序          | , ヿ , ー , ー , ー , ー , ー , ー , ー , 1 , 筆順序版本(CNS11643_1_98) |  |             |  |  |  |
| 漢字國屬性        | 台灣(TF-6C2D)  |  |             |  |  |  |

We found U+2386E on *Kangxi-Plus Unicode Han Dictionary*. Mr. Wáng and his team thought that it's the variant form of 擠 (U+3A71), but we couldn't find more information about this view.

Fig. 6.3 U+2386E on Kangxi-Plus Unicode Han Dictionary

张玉书, 陈廷敬, 王宏源 等 (Zhāng Yùshū, Chén Tíngjìng, Wáng Hóngyuán, et al): 《康熙字典(增订版)》(*Kangxi-Plus Unicode Han Dictionary*), 北京: 社会科学文献出版社 [Běijīng: Social Sciences Academic Press (CHINA)], 2015, ISBN 978-7-5097-7093-1, p. 579



We suggest three ways to solve this problem:

A1) Encode TF-6C2D as a new ideograph, and the kRSUnicode value for it is 75.21.

A2) Add a new U-source glyph into U+2386E, and change the kRSUnicode value for it of 64.21.

We don't think Fig. 6.3 has shown a favourable evidence for this glyph.

Table 6.1 Information for a New Proposed U-source Ideograph

| G       | lyph   | IDS                                   | Radical | Radical<br>Form | Strokes | First<br>Stroke | Simplified<br>Form | Variant<br>Form | Variant<br>Code Point | U-source<br>Number |
|---------|--------|---------------------------------------|---------|-----------------|---------|-----------------|--------------------|-----------------|-----------------------|--------------------|
| $\perp$ | 姓<br>討 | □ 才欝<br>U+2FF0,<br>U+624C,<br>U+23835 | 64      | 手               | 21      | 1               | N/A                | 灂               | U+3A71                | UCI-xxxxx          |

B1) Keep the present situation of U+2386E, and add a second kRSUnicode value of 64.21 in the future.

B2) Add two glyphs into SVS like this:

Table 6.2 Proposed SVSes for U + 2386E

| Base Character | Variation Selector | Glyph    | Description |
|----------------|--------------------|----------|-------------|
| U+2386E        | FE00 (VS1)         | <u> </u> |             |
| U+2386E        | FE01 (VS2)         | 機        | TF-6C2D     |

- C1) Keep the present situation of U+2386E, and add a second kRSUnicode value of 64.21 in the future.
- C2) Add two glyphs into IVD like this:

Table 6.3 Proposed IVSes for U+2386E

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+2386E        | E0100 (VS17)       | 攒     | UCS2003     |
| U+2386E        | E0101 (VS18)       | 欑     | TF-6C2D     |

## 7. Error T-source glyph for U+23A26

The IDS of the current UCS2003 glyph or G-sourc glyph for U+23A26 is  $\Box$  白田从戈, but the IDS of the current T-source glyph for U+23A26 is  $\Box$  白田 本文. (pls see Fig. 7.1) 日本文 shouldn't be unifiable with 日从戈.

Fig. 7.1 U+23A26 on The Unicode Standard, Version 9.0

We show the evidence in one of the original sources: GKX-0582.04 which is in *KangXi Dictionary*. (pls see Fig. 7.3)

Figs. 7.2 through 7.4 GKX-0582.04 Glyph on Different Versions of KangXi Dictionary

内府本《康熙字典》(午集備考), p. 1391



同文書局版《康熙字典》(午集備考), p. 582



张玉书, 陈廷敬, 王宏源 等 (Zhāng Yùshū, Chén Tíngjìng, Wáng Hóngyuán, et al): 《康熙字典(增订版)》(*Kangxi-Plus Unicode Han Dictionary*), 北京: 社会科学文献出版社 [Běijīng: Social Sciences Academic Press (CHINA)], 2015, ISBN 978-7-5097-7093-1, p. 738

**找 xiān\_**8. 12 集韻 殲27008古作茂。

We have found more evidences about this glyph.

## Fig. 7.5 GKX-0582.04 Glyph on Zhōnghuá Dà Zìdiǎn

徐元誥, 歐陽溥存, 汪長禄 (Xǔ Yuángào, Ōuyáng Pǔncún, Wāng Chánglù): 《中華大字典》 (*Zhōnghuá Dà Zìdiǎn*), 香港: 中華書局 [Hong Kong: Chung Hwa Book Co. (H.K.) Ltd.], 1958, 30686, p. 794



#### Fig. 7.6 GKX-0582.04 Glyph on Zhōngwén Dà Cìdiǎn

張其昀 等 (Zhāng Qíyún, et al.): 《中文大辭典》 (Zhōngwén Dà Cìdiǎn, aka Comprehensive Dictionary of the Chinese Language), 臺北: 中國文化研究所 (Taipei: Zhōngguó Wénhuà Yánjiūsuǒ), 1973, p. 7672



Fig. 7.7 GKX-0582.04 Glyph on Zhōnghuá Zìhǎi

冷玉龙, 韦一心 等 (Lěng Yùlóng, Wěi Yīxīn, et al): 《中华字海》 (*Zhōnghuá Zìhǎi*), 北京: 中华书局 (Běijīng: Zhonghua Book Company) & 北京: 中国友谊出版公司 (Běijīng: Friendship Publishing Corporation of China), 1994, ISBN 7-5057-0630-6/H•1, p. 810

**茂**同"奸"。见《直音篇》。

#### Fig. 7.8 GKX-0582.04 Glyph on Hànyǔ Dàzìdiǎn V2

汉语大字典编辑委员会(Hànyǔ Dàzìdiǎn Biānjí Wěiyuánhuì):《汉语大字典(第二版)》(*Hànyǔ Dàzìdiǎn V2*), 武汉: 湖北长江出版集团崇文书局 (Wǔhàn: Chongwen Publishing House of Hubei Changjiang Publishing Group) & 成都: 四川出版集团四川辞书出版社 (Chéngdū: Sichuan Reference Press of Sichuan Publishing Group), 2010, ISBN 978-7-5403-1744-7, p. 1512

## 片公 同"殲"。《集韻·鹽韻》:"殲,古作磁。"

The official website of the Master Ideographs Seeker for CNS11643 Chinese Standard Interchange Code shows this contradiction because of the disunification of two glyphs for U+23A26. (pls see Fig. 7.9)

Fig. 7.9 U + 23A26 on The Master Ideographs Seeker

| 全字庫字型                 | 發音                   |                                     |       |  |  |  |
|-----------------------|----------------------|-------------------------------------|-------|--|--|--|
| <b>茂</b> 茂 茂 明體 指體 宋體 |                      | 4 — 马<br>男 <b>聲€</b> ⊬女 <b>聲◆</b> € |       |  |  |  |
| 拼音                    | 漢語<br>jiān           |                                     |       |  |  |  |
| 字義                    | 筆畫                   | 部首                                  | 倉頡    |  |  |  |
| 相關詞                   | 13                   | 歹                                   | YBOOI |  |  |  |
| 部件                    | 上,口,一,体,戈            | ,部件版本(CNS11643                      | 2_94) |  |  |  |
| 筆順序                   | , 一 ,                |                                     |       |  |  |  |
| 漢字國屬性                 | 台灣(T6-5371) 中國(G_KX) |                                     |       |  |  |  |

The "MOE" Dictionary of Chinese Character Variants ("教育部" 異體字字典) shows the key

evidences for T6-5371. (pls see Fig. 7.10) When we review the key evidences such as Fig. 7.3, Fig. 7.6 and so on, we thought T6-5371 glyph was error for its evidences.

Fig. 7.10 U+23A26 on The "MOE" Dictionary of Chinese Character Variants



We suggest four ways to solve this problem:

- A) Ask T-source to change the T6-5371 glyph.
- B) Encode T6-5371 as a new ideograph, and the kRSUnicode value for it is 78.10.
- C1) Keep the present situation of U+23A26, and add a second kRSUnicode value of 78.10 in the future.
- C2) Add three glyphs into SVS like this:

Table 7.1 Proposed SVSes for U + 23A26

| Base Character     | Variation Selector | Glyph | Description |
|--------------------|--------------------|-------|-------------|
| U+23A26            | FE00 (VS1)         | 씑     | UCS2003     |
| U+23A26            | FE01 (VS2)         | 茂     | GKX-0582.04 |
| U+23A26 FE02 (VS3) |                    | 茂     | T6-5371     |

- D1) Keep the present situation of U+23A26, and add a second kRSUnicode value of 78.10 in the future.
- D2) Add three glyphs into IVD like this:

Table 7.2 Proposed IVSes for U+23A26

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+23A26        | E0100 (VS17)       | 茂     | UCS2003     |
| U+23A26        | E0101 (VS18)       | 茂     | GKX-0582.04 |
| U+23A26        | E0102 (VS19)       | 茂     | T6-5371     |

## 8. Error T-source glyph for U+24120

The IDS of the current UCS2003 glyph glyph for U+24120 is  $\square$   $\mathring{}$   $\overset{\circ}{\mathbb{Z}}$ , but the IDS of the current T-source glyph for U+24120 is  $\square$   $\mathring{}$   $\overset{\circ}{\mathbb{Z}}$ . (pls see Fig. 8.1)  $\overset{\circ}{\mathbb{Z}}$  shouldn't be unifiable with  $\overset{\circ}{\mathbb{Z}}$ .

Fig. 8.1 U+24120 on The Unicode Standard, Version 9.0

24120 水 85.18 UCS2003 TF-676A

The official website of the Master Ideographs Seeker for CNS11643 Chinese Standard Interchange Code shows this contradiction because of the disunification of two glyphs for U+24120. (pls see Fig. 8.2)

Fig. 8.2~U + 24120~on~The~Master~Ideographs~Seeker

| 全字庫字型 | 發音                                  |  |      |  |  |  |  |
|-------|-------------------------------------|--|------|--|--|--|--|
| 潜濱濱   | ター与 *<br>男 <b>聲◆</b> K女 <b>聲◆</b> K |  |      |  |  |  |  |
| 拼音    |                                     | 漢語<br>pín  |      |  |  |  |  |
| 字義    | 筆畫                                  | 部首   | 倉頡   |  |  |  |  |
| 相關詞   | 21                                  | 水  | ETJC |  |  |  |  |
| 部件    | 才,**,か,一,少,                         | 扌, 艹, 宀, 一, 少 , 貝 ,部件版本(CNS11643_2_94)                             |      |  |  |  |  |
| 筆順序   | 「つ, 一, 一, 一, 本(CNS11643_1_98)       | , ¬ , ー , ー , ー ,   , ノ , 、 , ー ,   , ¬ ,   , 筆順序版本(CNS11643_1_98) |      |  |  |  |  |
| 漢字國屬性 | 台灣(TF-676A)                         |  |      |  |  |  |  |

According to Fig. 8.2, we thought TF-676A glyph was error.

We suggest four ways to solve this problem:

- A) Ask T-source to change the TF-676A glyph for U + 24120.
- B1) Withdraw the TF-676A glyph for U + 24120.
- B2) Add a new U-source glyph into U+24120.

Table 8.1 Information for a New Proposed U-source Ideograph

| Glyph | IDS                                 | Radical | Radical<br>Form | Strokes | First<br>Stroke | Simplified<br>Form | Variant<br>Form | Variant<br>Code Point | U-source<br>Number |
|-------|-------------------------------------|---------|-----------------|---------|-----------------|--------------------|-----------------|-----------------------|--------------------|
| 濱     | □氵賞<br>U+2FF0,<br>U+6C35,<br>U+85B2 | 85      | Ÿ               | 18      | 1               | N/A                | N/A             | N/A                   | UCI-xxxxx          |

C1) Keep the present situation of U+24120, and add a second kRSUnicode value of 64.18 in the

future.

C2) Add two glyphs into SVS like this:

Table 8.2 Proposed SVSes for U+24120

| Base Character Variation Selector |            | Glyph | Description |
|-----------------------------------|------------|-------|-------------|
| U+24120                           | FE00 (VS1) | 濱     | UCS2003     |
| U+24120                           | FE01 (VS2) | 攢     | TF-676A     |

- D1) Keep the present situation of U+24120, and add a second kRSUnicode value of 64.18 in the future.
- D2) Add two glyphs into IVD like this:

Table 8.3 Proposed IVSes for U+24120

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+24120        | E0100 (VS17)       | 濱     | UCS2003     |
| U+24120        | E0101 (VS18)       | 攢     | TF-676A     |

## 9. Wrong V-source glyph design for U+244C7

The IDS of the current UCS2003 glyph glyph for U+244C7 is ①火瓊, but the IDS of the current T-source glyph for U+244C7 is ②火玉目日夕因目久. (pls see Fig. 9.1)

Fig. 9.1 U+244C7 on The Unicode Standard, Version 9.0



The offical website of VNPF shows a new V-source glyph for U+244C7.

Fig. 9.2 *U* + 244C7 on *VNPF* 

| Quốc Ngữ | Hán-Nôm | Codepoint      | Context    | Ref. | English |
|----------|---------|----------------|------------|------|---------|
| quầng    | 爈       | <u>U+244c7</u> | quầng sáng | vhn  | aureola |

This mistake had been pointed out in IRG #36, and V-source answered that the V-glyph would be updated. (pls see <a href="IRGN1792Updated V Font for CJK\_B">IRGN1792Updated V Font for CJK\_B</a> or Fig. 9.3)

Fig. 9.3 IRGN1792Updated V Font for CJK\_B, p. 1

| UCS Code | Glyph in<br>UCS2003 | Glyph in IRG N1748                    | Updated font |
|----------|---------------------|---------------------------------------|--------------|
| 23E54    | 迎<br>UCS2003        | 23E54<br>* 85.9                       | 迥            |
| 24448    | 燈 UCS2003           | 24448<br>火 86.14                      | 焇敳           |
| 244C7    | 煌。                  | 244C7<br>火 86.18 炬<br>UCS2003 V3-337D | 熳            |

We suggest that V-source update V-glyph like the glyph on VNPF as soon as possible.

#### 10. Wrong G-source glyph design for U + 2459B

Fig. 10.1 U + 244C7 on The Unicode Standard, Version 9.0

2459B 月 90.12 UCS2003 GHZ-42378.21 T5-5544

When we reviewed one of the original sources: GHZ-42378.21, we thought the G-source glyph design was wrong. (pls see Fig. 10.2) The IDS of GHZ-42378.21 glyph should be □將且.

Fig. 10.2 GHZ-42378.21 on Hànyǔ Dàzìdiǎn V1

上達 jiǎng《集韻》子朝切,上蕩精。又《字彙補》子兩切。 大。《類篇・爿部》:"蜚,大也。"

Fig. 10.3 IRGN2179Unified&WithdrawnV3.0, p. 25

| 1 |       | 11 | 1 7 | I. | 1 |           |                   |
|---|-------|----|-----|----|---|-----------|-------------------|
|   |       | 90 |     |    |   | 出会        |                   |
|   | 02343 | Ħ  |     |    |   |           | withdrawn, irg46. |
|   |       | 12 | 3   |    |   | UTC-02856 |                   |
| T |       |    |     |    |   | 6         |                   |

Fig. 10.4 UTC-02856 on BabelStone

| U+EB49 | ূ | ⊟將且 | U+2C312 0 | HFDC p. 6956 | UTC-02856 | Ext. G<br>Candidate<br>(IRGN2107) |
|--------|---|-----|-----------|--------------|-----------|-----------------------------------|
|--------|---|-----|-----------|--------------|-----------|-----------------------------------|

We thought UTC-02856 should be unifiable with U + 2459B, so we suggest two ways to solve this problem:

- A) Ask G-source to change the GHZ-42378.21 glyph for U + 2459B.
- B1) Withdraw the GHZ-42378.21 glyph for U + 2459B.
- B2) Make UTC-02856 be unifiale with U + 2459B.

## 11. Disunification for U+24662

Fig. 11.1 U + 24662 on The Unicode Standard, Version 9.0



We show the evidence in one of the original sources: GKX-0701.13 which is in *KangXi Dictionary*. (pls see Fig. 11.3)

Figs. 11.2 through 11.4 GKX-0701.13 on different versions of KangXi Dictionary 内府本《康熙字典》, p. 1678





张玉书, 陈廷敬, 王宏源 等 (Zhāng Yùshū, Chén Tíngjìng, Wáng Hóngyuán, et al): 《康熙字典(增订版)》(*Kangxi-Plus Unicode Han Dictionary*), 北京: 社会科学文献出版社 [Běijīng: Social Sciences Academic Press (CHINA)], 2015, ISBN 978-7-5097-7093-1, p. 905

According to Fig. 11.5, we thought T5-334B was a different character from GKX-0701.13.

Fig. 11.5 U + 24662 on The Master Ideographs Seeker

| 全字庫字型             | 發音                                     |   |       |  |  |  |  |
|-------------------|--|---|-------|--|--|--|--|
| 特 祥 特<br>明體 格體 宋體 | 口`<br>男聲 <mark>《</mark> 於女 <b>聲《</b> 於 |   |       |  |  |  |  |
| 拼音                |  | 漢語<br>yù  |       |  |  |  |  |
| 字義                | 筆畫                                     | 部首  | 倉頡    |  |  |  |  |
| 相關詞               | 11                                     | 牛   | HQTMD |  |  |  |  |
| 部 件               | 华,→+,于,部件#                             | 华,→+,于,部件版本(CNS11643_2_94)  |       |  |  |  |  |
| 筆順序               | /,-, ,-,<br>(CNS11643_1_94)            | ノ, 一, 丨, 一, 一, 一, 一, 一, 一, <sup>1</sup> , <sub>筆順序版本</sub><br>(CNS11643_1_94) |       |  |  |  |  |
| 漢字國屬性             | 台灣(T5-334B) 中國                         | 図(G_KX)   |       |  |  |  |  |

We suggest three ways to solve this problem:

- A) Encode T5-334B as a new ideograph, and the kRSUnicode value for it is 93.7.
- B) Keep the present situation of U + 24662, and add three glyphs into SVS like this:

Table 11.1 Proposed SVSes for U + 24662

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+24662        | FE00 (VS1)         | 桦     | UCS2003     |
| U+24662        | FE01 (VS2)         | 牪     | GKX-0701.13 |
| U+24662        | FE02 (VS3)         | 牸     | T5-334B     |

C) Keep the present situation of U+24662, and add three glyphs into IVD like this:

Table 11.2 Proposed IVSes for U+24662

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+24662        | E0100 (VS17)       | 牪     | UCS2003     |
| U+24662        | E0101 (VS18)       | 牪     | GKX-0701.13 |
| U+24662        | E0102 (VS19)       | 特     | T5-334B     |

## 12. Error T-source glyph for U+26C16

The IDS of the current UCS2003 glyph or G-source glyph for U+26C16 is = +  $\pm$   $\pm$ , but the IDS of the current T-source glyph for U+26C16 is =  $\pm$   $\pm$   $\pm$   $\pm$   $\pm$   $\pm$  shouldn't be unifiable with  $\pm$ .

Fig. 12.1 U+26C16 on The Unicode Standard, Version 9.0

26C16 艸 140.8

UCS2003 GHZ-53214.11 T5-3B5D

The official website of the Master Ideographs Seeker for CNS11643 Chinese Standard Interchange Code shows this contradiction because of the disunification of two glyphs for U+26C16. (pls see Fig. 12.2)

Fig. 12.2 U + 26C16 on The Master Ideographs Seeker

| 全字庫字型          | 發音  |    |      |
|----------------|---|----|------|
| 趀 趀 箍 明體 楷體 宋體 | メモ゛<br>男撃 <mark>・</mark> (大 <mark>撃・(</mark> ): |    |      |
| 拼音             | 漢語<br>wò  |    |      |
| 字義             | 筆畫  | 部首 | 倉頡   |
| 相關詞            | 12  | 艸  | TQKS |
| 部件             | 体 , 扌 , 九 , カ , 部件版本(CNS11643_2_94)             |    |      |
| 筆順序            | , 一, 一 , 一 , 一 , 一 , 一 , 一 , 一 , 1 , 1          |    |      |
| 漢字國屬性          | 台灣(T5-3B5D) 中國(G_HZ) 北韓(KP1-6E6D)               |    |      |

We suggest three ways to solve this problem:

- A) Encode T5-3B5D as a new ideograph, and the kRSUnicode value for it is 118.8.
- B1) Keep the present situation of U+26C16, and add a second kRSUnicode value of 118.8 in the future.
- B2) Add three glyphs into SVS like this:

Table 12.1 Proposed SVSes for U + 26C16

| Base Character | Variation Selector | Glyph | Description  |
|----------------|--------------------|-------|--------------|
| U+26C16        | FE00 (VS1)         | 菰     | UCS2003      |
| U+26C16        | FE01 (VS2)         | 菰     | GHZ-53214.11 |
| U+26C16        | FE02 (VS3)         | 箍     | T5-3B5D      |

- C1) Keep the present situation of U+26C16, and add a second kRSUnicode value of 118.8 in the future.
- C2) Add three glyphs into IVD like this:

Table 12.2 Proposed IVSes for U+26C16

| Base Character | Variation Selector | Glyph | Description  |
|----------------|--------------------|-------|--------------|
| U+26C16        | E0100 (VS17)       | 菰     | UCS2003      |
| U+26C16        | E0101 (VS18)       | 菰     | GHZ-53214.11 |
| U+26C16        | E0102 (VS19)       | 箍     | T5-3B5D      |

## 13. Error T-source glyph for U+281AF

The IDS of the current UCS2003 glyph or G-source glyph or H-source for U+281AF is  $\square$ 足僕, but the IDS of the current T-source glyph for U+281AF is  $\square$ 足業. (pls see Fig. 13.1) 菐 shouldn't be unifiable with 僕.

Fig. 13.1 U+281AF on The Unicode Standard, Version 9.0

281AF 足 157.14 UCS2003 GKX-1235.21 T7-5544 比美 H-9D41

The official website of the Master Ideographs Seeker for CNS11643 Chinese Standard Interchange Code shows this contradiction because of the disunification of two glyphs for U+281AF. (pls see Fig. 13.2) And T7-5544 should be unifiable with U+8E7C.

Fig. 13.2 U + 281AF on The Master Ideographs Seeker

| 全字庫字型 | 發音   |   |    |  |
|-------|--|---|----|--|
| 跳跳 跳  | タメ <sup>*</sup><br>男 <b>聲『</b> ド女 <b>聲』</b> [ド |   |    |  |
| 拼音    |  | 漢語<br>pú                                    |    |  |
| 字義    | 筆畫   | 部首  | 倉頡 |  |
| 相關詞   | 21   | 21 足 RMOTO                                  |    |  |
| 部件    | 口,此,此,夫  | 口,止,些, ** , ** , ** , ** , ** , ** , ** , * |    |  |
| 筆順序   | , つ , 一 , 一 , 一 , 一 , 一 , 一 ,                  |   |    |  |
| 漢字國屬性 | 台灣(T7-5544) 中國(G_KX) 北韓(KP1-7BAF)              |   |    |  |

We suggest three ways to solve this problem:

- A) Make T7-5544 be unifiable with U + 8E7C.
- B1) Keep the present situation of U+281AF, and add a second kRSUnicode value of 157.12 in the future.
- B2) Add four glyphs into SVS like this:

Table 13.1 Proposed SVSes for U + 281AF

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+281AF        | FE00 (VS1)         | 蹼     | UCS2003     |
| U+281AF        | FE01 (VS2)         | 蹼     | GKX-1235.21 |
| U+281AF        | FE02 (VS3)         | 蹼     | T7-5544     |
| U+281AF        | FE03 (VS4)         | 蹼     | H-9D41      |

- C1) Keep the present situation of U+281AF, and add a second kRSUnicode value of 157.12 in the future.
- C2) Add four glyphs into IVD like this:

Table 13.2 Proposed IVSes for U+26C16

| Base Character | Variation Selector | Glyph | Description |
|----------------|--------------------|-------|-------------|
| U+281AF        | E0100 (VS17)       | 蹼     | UCS2003     |
| U+281AF        | E0101 (VS18)       | 蹼     | GKX-1235.21 |
| U+281AF        | E0102 (VS19)       | 蹼     | T7-5544     |
| U+281AF        | E0103 (VS20)       | 蹼     | H-9D41      |

## 14. Proposed additional information on CJK Unified Ideographs to ISO/IEC 10646

## Annex P

At last, we suggest the twelve error information could be included to ISO/IEC 10646 Annex P before they will be solved. One error information had been included in Annex P.

Table 14.1 Proposed Additional Information on CJK Unified Ideographs to ISO/IEC 10646

#### Annex P

| UCS / Glyph   | Additional Information                                  |
|---|---|
| 21846 如 加   | V0-3521 source glyph for this code point was mistakenly |
| UCS2003 V0-3521   | designed.   |
| 21BC0<br>位 42.29<br>图 图 图 图 图 图 图 图 图 图 图 图 图 图 图 图 图 图 图 | V2-7456 source glyph for this code point was mistakenly |
| UCS2003 V2-7456   | designed.   |
| 2278B <b>倭 倭</b>  | GHZ-10207.12 source glyph for this code point was       |
| UCS2003 GHZ-10207.12                                      | mistakenly designed.                                    |
| 22E3F 大佐<br>= 64.16 十年 大佐                                 | TF-6246 source glyph for this code point was mistakenly |
| UCS2003 TF-6246   | designed.   |
| 2386E + ** /5.21 ** ** ** ** ** ** ** ** ** ** ** ** **   | UCS2003 glyph for this code point was mistakenly        |
| UCS2003 TF-6C2D   | designed.   |
| 23A26 古公 古公 古谷  | T6-5371 source glyph for this code point was mistakenly |
| UCS2003 GKX-0582.04 T6-5371                               | designed.   |
| 24120 潜 增   | TF-676A source glyph for this code point was mistakenly |
| UCS2003 TF-676A   | designed.   |

| 244C7                        | T3-337D source glyph for this code point was mistakenly |
|------------------------------|---|
| UCS2003 V3-337D              | designed.   |
| 2459B                        | UCS2003 and GHZ-42378.21 source glyph for this code     |
| UCS2003 GHZ-42378.21 T5-5544 | point was mistakenly designed.                          |
| 24662                        | The T5-334B source glyph should have been unified with  |
| UCS2003 GKX-0701.13 T5-334B  | U+24662.  |
| 26C16 茄 茄 紡                  | T5-3B5D source glyph for this code point was mistakenly |
| UCS2003 GHZ-53214.11 T5-3B5D | designed.   |
| 281AF                        | T7-5544 source glyph for this code point was mistakenly |
| H-9D41                       | designed.   |