

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2/IRGN2149

Date: 2016-05-25

Source:	CJK editorial group
Title:	Editorial Report on CJK_F and Miscellaneous Issue
Meeting:	IRG#46, Beijing, China
Status :	To be Approved by IRG plenary
Actions required	To be followed by IRG members
Distribution:	IRG
Medium :	Electronic
Page:	2
Appendix:	1
Reference:	IRGN2146 and response, IRGN2152, IRGN2019 and response

Editors from China, Japan, R.O.Korea, TCA, met during IRG#46. Mr. Michel Suignard (UCS Project Editor) and representatives of SAT attended the meeting. The editorial group discussed the Following issues.

1. CJK_F ballot Comments

Reference:

- IRGN2146 Draft disposition of comments on ISO/IEC CD.2 10646 5th edition
 - China Responses
 - Japan Responses
 - SAT Responses
 - UK Responses
 - Feedbacks to UK (from Mr. Suzuki)
- IRGN2152 Disposition of comments on ISO/IEC CD.2 10646 5th edition (Updated IRGN2146 according to CJK_F editorial meeting during IRG#46)
 - Japan Feedback

The editors reviewed and discussed all responses.

The editors confirmed to accept the proposed dispositions by the UCS Project Editor except for the Japan ballot comments on unification.

The Japan ballot comments on unification and their corresponding responses were specially discussed, some characters removed from previous CJK_F internal lists will be restored. For important discussions and conclusions on these comments, see

appendix.

The editors then confirmed IRGN2152 which was generated according to conclusions, by the UCS Project Editor.

The CJK_F related parties submitted the updated fonts to the UCS Project Editor by 2016-05-25. The UCS Project Editor will generate new CJK_F Code Charts according to the conclusions.

2. IRG PnP

Reference:

- IRGN2092 IRG Principles and Procedures(IRG PnP) Version 8 draft Version3
 - Mr. Suzuki's Feedback on Source Reference
 - Japan's Feedback on Stroke Count

The editors accepted Mr. Suzuki's suggestion that the character naming rules should be explicitly specified in IRG PnP, in order to keep consistency between IRG and WG2.

The editors accepted Japan's suggestions on revisions on stroke count specified in IRG PnP.

The Rapporteur agrees to revise IRG PnP and supply a revised draft by 2016-07-15 for review by IRG. The revised version will be confirmed in next IRG meeting.

3. Horizontal Extension of Vietnamese Characters

Reference: I

- IRGN2132 Horizontal Extension of Vietnamese Characters
 - Responses from China, Japan, ROK and TCA

Editors from China, Japan, ROK and TCA reviewed IRGN2132 and provided feedback for the action by Vietnam. Vietnam is asked to follow up.

IRGN2146: Discussion Record on CJK F Unifications

suzuki toshiya, 2016-05-25

2	2D0D6 <small>刀 18.6</small> 列 <small>GZ-4861102</small>	52BD <small>力 19.6</small> 列 <small>GE-2353</small>	列 <small>T4-2550</small>	列 <small>J1-3384</small>	列 <small>K2-2532</small>
---	--	--	-----------------------------	-----------------------------	-----------------------------

2D0D6 and 52BD

Do not unify as they are not cognate and have different radicals.

if non-cognate, they should be disunified.

3	2D0E9 <small>刀 18.8</small> 𠄎 <small>GZ-4062201</small>	2B9AF <small>力 19.8</small> 𠄎 <small>GZFY-00521</small>
---	--	--


2D0E9 and 2B9AF

Do not unify as they are not cognate and have different radicals.

if non-cognate, they should be disunified.

Result: UK comments are approved. they are disunified because non-cognate.

7	2D421 爰 35.5 USAT-02739	225BB 心 61.3 UCS2003 GKK-0377.07 T6-275C
---	-------------------------------	--

2D421 and 225BB 

The components 心 and 必 have different meanings and are not unifiable. Therefore 2D421 should not be unified with 225BB.


if disunified by shape difference, JMJ-058052 and JMJ-059628 should be backed to Ext F

璦 JMJ-058052 04639 JMJ-058052	74A6 璦 96.13 璦 璦 璦 璦 璦 G1-6828 HB1-C0F5 T1-7440 J1-4C5D K2-484F U+74A6	unified by U+74A6, irg40.
---	--	---------------------------

爰 JMJ-059628 01644 JMJ-059628	611B 心 61.9 爰 爰 爰 爰 爰 G1-302E HB1-B752 T1-6378 J0-3026 K0-6471 U+0611B	unified by U+0611B, irg40.
---	--	----------------------------

Result: UK comments are approved. 2 JMJ characters would be restored in CJK F.

8	2D44B 夫 37.8 矩 KC-00818	77E9 矢 111.5 矩 矩 矩 矩 矩 矩 G0-3E58 HB1-AF78 T1-567B J0-366B K0-4F3B V1-605D
19	2DB0F 木 75.12 榧 KC-07047	69FB 木 75.11 榧 榧 榧 榧 榧 G3-416A HB2-E5E2 T2-512B J0-4450 K0-5027
20	2DB54 欠 76.8 欸 JMJ-057882	238C7 欠 76.7 欸 欸 UCS2003 GHZ-80027.22


2D44B and 77E9 
2DB0F and 69FB
2DB54 and 238C7

The components 夫 and 矢 have different meanings and are not unifiable. Therefore 2D44B should not be unified with 77E9, 2DB0F should not be unified with 69FB, and 2DB54 should not be unified with 238C7.

agreed, however 夫 and 矢 meaning difference is not valid objection. 器 should be disunified because 大 and 犬 are different meanings?

Result: UK comments are approved. no change in CJK F.

12	2D81D 手 64.5 USAT-05471	拘	22AE4 手 64.6 UCS2003 GKX-0428.07 T6-3152	拘 拘 拘
47	2E82F 酉 164.5 KC-05802	酌	2882E 酉 164.6 UCS2003 GHZ-63581.11 T6-594D	酌 酌 酌

2D81D and 22AE4 
 2E82F and 2882E


The components 𠂇 and 𠂈 have different meanings and are not unifiable. Therefore
 if disunified by shape, USAT-01643 and USAT01911 should be backed to Ext F

胸 Z_SAT01643 05972 USAT01643	80F8 肉 130.6 G0-5058 HB1-AFDD T1-5761 J0-363B K0-7D58 V1-637B J+80F8	胸 胸 胸 胸 胸 胸	unified by u+80f8, irg40. 
胷 Z_SAT01911 05973 USAT01911	80F7 肉 130.6 GE-3958 J+80F7	胷 胷 胷	unified by u+80f7, irg40.

Result: UK comments are approved. 2 SAT characters would be restored in CJK F.

16	2DA99 木 75.7 𣎵 JMJ-057740	6840 木 75.6 𣎵 𣎵 𣎵 𣎵 𣎵 𣎵 G0-686E H81-AEE5 T1-5629 J0-5B5C K0-4B7A V1-5971
----	------------------------------------	---

32	2E1D3 米 119.5 𣎵 JMJ-058287	25E4F 米 119.5 𣎵 𣎵 𣎵 UCS2003 GKX-0908.12 T6-4C63
----	-------------------------------------	--

2DA99 and 6840 
 2E1D3 and 25E4F


The components 𣎵 and 夕 have different meanings and are not unifiable. Therefore 2DA99 should not be unified with 6840, and 2E1D3 should not be unified with 25E4F.

if disunified by shape, JMJ-057872 should be backed to Ext F

𣎵 JMJ-057872 03695 JMJ-057872	2C155 木 75.20 𣎵 TE-4826 U+2C155	unified by U+2C155(CJK_E), irg43.
--	---	-----------------------------------

Result: UK comments were not approved. JMJ-057740 and JMJ-058287 are removed from CJK F. No restoration in this case.

27	2DEF3 田 102.12 疊 JMJ-058106	2DEF8 田 102.14 疊 JMJ-058108
33	2E232 糸 120.9 緋 USAT-01467	27322 虫 142.8 繪 繪 繪 UCS2003 GKX-1088.22 T5-4961
46	2E769 車 159.9 輻 JMJ-060221	2833A 車 159.8 翰 翰 UCS2003 H-8ECA

2DEF3 and 2DEF8 
 2E232 and 27322
 2E769 and 2833A


The components 人 and 口 have different meanings and are not unifiable. Therefore 2DEF3 should not be unified with 2DEF8, 2E232 should not be unified with 27322, and 2E769 should not be unified with 2833A.

if disunified by shape, JMJ-057168 should be backed to Ext F

喪 JMJ-057168 01028 JMJ-057168	55AA 口 30.9 喪 喪 喪 喪 喪 喪 G1-4925 HB1-B3E0 T1-5E46 J0-4153 K0-5F43 V1-4F64 U+55AA	unified to U+55AA, irg42. postponed being unified by U+55AA, irg40.
--	---	---

Result: UK comments are approved. 1 JMJ characters would be restored in CJK F.

28	2E0DD <small>禾 115.9</small>  種 <small>USAT-02356</small>	2581E <small>禾 115.9</small> 穉 穉 <small>UCS2003 GHZ-42617.16</small>
----	---	--

2E0DD and 2581E 

The right hand components of 2E0DD and 2581E are quite different, so 2E0DD should not be unified with 2581E.

if disunified by shape, JMJ-060109 should be backed to Ext F

蕊 <small>JMJ-060109</small> 06182 <small>JMJ-060109</small>	26BCC <small>艸 140.7</small> 蕊 蕊 蕊 <small>UCS2003 GKX-1035.05 T6-4438</small> 	unified by U+26BCC, irg43. 
---	--	---

Result: UK comment was approved by the structural difference, not the component. No restoration for this case.

1	2D0D2 刀 18.5 刳 USAT-00085	206B9 刀 18.4 刳 刳 刳 UCS2003 GKX-0137.08 T5-222E
---	------------------------------------	---

10	2D6E1 彳 60.5 役 JMJ-057485	5F79 彳 60.4 役 役 役 役 役 役 G0-525B HB1-A7D0 T1-4A32 J0-4C72 K0-6635 V1-545C
----	------------------------------------	---

11	2D818 手 64.5 投 投 JMJ-057554 USAT-03285	6295 手 64.4 投 投 投 投 投 投 G0-4D36 HB1-A7EB T1-4A4D J0-456A K0-7761 V1-565A
----	---	---

22	2DB90 歹 78.5 殳 JMJ-057905	6B81 歹 78.4 殳 殳 殳 殳 G0-6962 H-8DD1 T3-2958 J4-2F7E
----	------------------------------------	---


23	2DBB8 攴 79.7 𠂔 JMJ-057919	3C7F 攴 79.6 𠂔 𠂔 G5-5063 T3-316B
----	------------------------------------	--

24	2DBC0 攴 79.10 殿 JMJ-057926	6BBF 攴 79.9 殿 殿 殿 殿 殿 殿 G0-356E HB1-B7B5 T1-645B J0-4542 K0-6E7C V1-5B48
----	-------------------------------------	---

25	2DBC2 攴 79.11 𠂔 JMJ-057928	6BC4 攴 79.10 𠂔 𠂔 𠂔 𠂔 𠂔 GE-2F64 HB2-E26A T2-4B32 J1-4635 K1-5830
----	-------------------------------------	--

35	2E304 耳 128.12 聲 JMJ-058370	8072 耳 128.11 聲 聲 聲 聲 聲 聲 G1-4979 HB1-C16E T1-7478 J0-6661 K0-6122 V1-6363
----	--------------------------------------	---

45	2E752 車 159.5 輟 JMJ-058827	8ED7 車 159.4 輟 輟 輟 GE-3F62 HB2-D7DF T2-3962
----	-------------------------------------	--

2D0D2 and 206B9
 2D6E1 and 5F79 
 2D818 and 6295
 2DB90 and 6B81
 2DBB8 and 3C7F
 2DBC0 and 6BBF
 2DBC2 and 6BC4
 2E304 and 8072
 2E752 and 8ED7

The components 𠄎 and 𠄎 are variants, but the difference in shape is significant, and they should not be unifiable. Therefore do not unify these nine pairs of characters.


if disunified by shape, JMJ-056916, JMJ-057920, JMJ-057921, JMJ-058430, JMJ-058742 should be backed to Ext F

𠄎 JMJ-056916 00250 JMJ-056916	4F07 人 94 𠄎 𠄎 G5-315B T3-2274 U+4F07	unified by U+4F07, irg40.
般 JMJ-057920 03827 JMJ-057920	6BB7 𠄎 79.6 般 般 般 般 般 般 G0-5273 HB1-AEEF T1-5633 J0-5D56 K0-6B5C V1-5B45 U+6BB7	unified by U+6BB7, postponed changing radical to 𠄎(weapon), irg40. irg40.
𠄎 JMJ-057921 03831 JMJ-057921	6BB8 𠄎 79.7 𠄎 𠄎 𠄎 G5-5065 T4-324E J1-4630 U+6BB8	unified by U+6BB8, postponed changing radical to 𠄎(weapon), irg40.
般 JMJ-058430 06106 JMJ-058430	822C 舟 137.4 般 般 般 般 般 般 G0-3063 HB1-AFEB T1-576F J0-484C K0-5A75 V1-644B U+822C	unified to U+822C, irg41.
設 JMJ-058742 06826 JMJ-058742	8A2D 言 149.4 設 設 設 設 設 設 G1-4968 HB1-B35D T1-5D43 J0-405F K0-6062 V1-667D U+8A2D	unified to U+8A2D, irg41.

Result: UK comments were approved. 5 JMJ characters would be restored to CJK F.

26	2DBF6 氏 83.4 岷 JMJ-057936	6C13 氏 83.4 岷 岷 岷 岷 岷 岷 G0-4325 HB1-AA5D T1-4E40 J0-5D62 K0-586C V1-5B58
34	2E295 网 122.9 罔 JMJ-058343	262DF 网 122.9 罔 罔 罔 UCS2003 GKX-0948.36 T7-2675
37	2E330 肉 130.10 羸 JMJ-058394	267C0 肉 130.10 羸 羸 羸 UCS2003 GHZ-32104.04 T6-6070

2DBF6 and 6C13

2E295 and 262DF 

2E330 and 267C0

The components 凵 and 亡 are variants, but the difference in shape is significant, and they should not be unifiable. Therefore do not unify these three pairs of characters.


if disunified by shape, JMJ-058447 should be backed to Ext F

荒 JMJ-058447 06156 JMJ-058447	8352 艸 140.6 荒 荒 荒 荒 荒 荒 G0-3B44 HB1-AFEE T1-5772 J0-3953 K0-7C58 V1-647E U+8352	unified to U+8352, irg41.
--	--	---------------------------

Result: UK comments were approved. 1 JMJ character would be restored in CJK F.

31	2E1CF 米 119.4 USAT-03923	杭	7C87 米 119.4 GE-375D	杭	杭	杭	杭	T3-3330 J1-532D K2-5146
----	--------------------------------	---	----------------------------	---	---	---	---	-------------------------------

48	2E9E0 革 177.4 USAT-03573	𪗇	29273 革 177.4 UCS2003	𪗇	𪗇	GHZ-74327.20
----	--------------------------------	---	-----------------------------	---	---	--------------

2E1CF and 7C87 
 2E9E0 and 29273


In many cases 𪗇 is a variant of 𪗇, but the difference is quite significant, and it is not clear

if disunified by shape, JMJ-006465 should be backed to Ext F

𪗇 JMJ-006465 00492 JMJ-006465	4EA2 𪗇 𪗇 𪗇 𪗇 𪗇 𪗇 <small>G0-3F3A HB1-A4AE T1-446D J0-5036 K0-7971 V1-4A5B</small> U+4EA2	unified to U+4EA2, but not agreed to be a general example, change to R16.0(𪗇), SC=2 and FS=3, irg42. postponed being unified by U+4EA2, postponed changing Radical to 𪗇(table), irg40.
--	--	--

Result: UK comments are not approved. 2 SAT characters are removed from CJK F. No restoration for this case.

41	2E484 褱 <small>紳 140.13</small> USAT-03743	2E5AB 褱 <small>衣 145.10</small> KC-06763
----	---	---

2E484 and 2E5AB 


There is no precedence for unifying 壯 and 壯, which in China are treated as traditional and

According to SCS, they should be unified

壯 壯 GTJ
 58EE 58EF

Result: UK comment was not approved. 2 CJK F characters are unified.

42	2E5E8 見 147.14 𧄀 JMJ-058726	2789B 見 147.14 𧄀 𧄀 UCS2003 GHZ-63676.07
----	--------------------------------------	--

2E5E8 and 2789B 

It is not clear that these two characters are cognate, and should be unified. 2E484 consists

if disunified by shape, USAT-060197, KA-KC07005, JMJ-057036 should be backed to Ext F

𧄀 Z_SAT60197 00465 USAT60197	712D 火 86.8 𧄀 𧄀 𧄀 G5-3344 T3-3D21 J1-496D K1-5852	unified to U+712D, irg41.
𧄀 KA-KC07005 01989 KC-07005	21C08 儿 43.10 𧄀 𧄀 UCS2003 GHZ-80012.22	unified to U+21C08, irg41. radical to 儿 (legs) U+2F09, irg40.
兜 JMJ-057036 00630 JMJ-057036	515C 儿 10.9 兜 兜 兜 兜 兜 兜 G0-3635 HB1-B0C2 T1-5927 J0-3375 K0-545F V1-4C42	unified to U+515C, irg42. postponed being unified by U+515C, but sc12, fs3, irg40.

Result: UK comment was not approved, JMJ-058726 is removed from CJK F. No restoration for this case.

Disunification Proposed by UK (T6)

2D0F0 剡 剡
刀 18.8
JMJ-059378 USAT-04376

The number of the characters including (U+21585) is 13, that of the characters including 死 is 73. The existing disunification are 2:

- 1) 死(U+21585) 死(U+6B7B),
- 2) 葬(U+2BE18) 葬(U+22348).

In CJK F, following 2 might be similar cases.

2D726 怨
心 61.5
JMJ-057499 =怨

2D729 𢦏
心 61.6
USAT-60105 =怨

2D4E5 宛
宀 40.5
JMJ-057329 =亥

2D4E9 宛 宛
宀 40.6
JMJ-057331 USAT-04051 =亥(J), 完(SAT)

Result: UK Comment was not approved. JMJ-059378 and USAT-04376 are kept as unified.

JMJ-057499 and USAT-60105 was not unified because too significant difference by the top horizontal stroke.

JMJ-057329, JMJ-057331 and USAT-04051 are regarded as unifiable. Among 2 JMJ characters, JMJ-057329 was chosen to be kept in CJK F code chart.

Disunification proposed by UK (T9)

2DDC6

卅 90.4

收

妝

JMJ-057583

USAT-04734

T9. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2DDC6

U+2DDC6. The two source glyphs (JMJ-057583 and USAT-04734) have different left hand components, which are not unifiable. Moreover, the two characters are not cognate: JMJ-057583 is a variant of U+6536 收 (See Zhonghua Zihai p. 1006); whereas USAT-04734 is a variant of U+7267 妝 (See Zhonghua Zihai p. 1006).

Proposed change by U.K.

Encode JMJ-057583 and USAT-04734 as separate characters.

Result: Although no consensus is formed to handle the comments based on Zhonghua Zihai, China and TCA mentioned their future possibilities to standardize USAT-04734 glyph as non-cognate characters. Considering their possibilities, SAT suggested to disunify them. Although this pair is disunified in CJK F, the pairs including these characters as the components should be unified.

Disunification Proposed by UK (T13)



Following pairs might be considered for the consistency.



Result: UK comment was not approved. USAT-00811 and USAT-02508 are unified (USAT-02508 would be kept).

JMJ-057942 is non-cognate with U+6C97, it would be kept in CJK F.

JMJ-057747 is non-cognate with U+23440, it would be kept in CJK F.

JMJ-057640 is cognate with U+2B780, it would be removed from CJK F.