

**ISO/IEC JTC1/SC2/WG2
Coded Character Set
Secretariat: Japan (JISC)**

Doc. Type: Draft disposition of comments

Title: Draft disposition of comments on ISO/IEC CD.2 10646 5th edition

Source: Michel Suignard (project editor)

Project: JTC1.02.10646.00.00.00.05

Status: For review by WG2

Date: 2016-04-23

Distribution: WG2

Reference: SC2 N4446 N4454, various IRG documents

Medium: Paper, PDF file

Comments were received from China, India, Japan, Mongolia, Norway, UK, and USA. The following document is the disposition of those comments. The disposition is organized per country.

Note – With some minor exceptions, the full content of the ballot comments have been included in this document to facilitate the reading. The dispositions are inserted in between these comments and are marked in **Underlined Bold Serif text**, with explanatory text in *italicized serif*.

The comments received can be disposed in a way to create consensus for most of the topics:

- Mongolian: move back the code chart to only show standardized variants that were agreed in the previous edition, and remove any annotation that were added since that former edition (presentation forms stay)
- Miao: fix code positions as requested by China and US
- Tamil: remove the new Tamil Supplement block and related references in existing blocks
- Nushu: re-order 10 characters as requested by UK
- Various minor editorial issues

This leaves CJK Ideographic matters as requiring substantial work, especially CJK Unified Ideographs Extension F. Therefore, in consultation with CJK experts, it was decided to wait for the next IRG meeting (#46 to be held in Beijing, China, May 23-27) to go over relevant comments and proposed disposition and create a consensus position allowing to generate an enquiry ballot (DIS) as a result.

China: Negative

Technical comments on Mongolian

(The same comments were made by Mongolia)

T1. Principle errors

First, this scheme seriously disrupts the more than 8 hundred years of formation and solidity of the complete traditional Mongolian orthography. It needs to fully and correctly reflect the diversity of traditional Mongolian orthography of the Mongolian character encoding and to be in line with the spirit of global protection and development of culture diversity. The information technology shall adapt to and solve the particular process of the traditional Mongolian orthography and must not be adapted to by breaking the Mongolian orthography for the limitation of words processing function of some current computer systems! This is the relationship between purpose and tools, must not put the cart before the horse.

Second, with the 8 hundred years of history of the development of traditional Mongolian script, its orthography has been enriched and improved. There are orthographic differences reflected in the written Mongolian sources of different times. Traditional Mongolian character encoding should include and reflect the whole process of traditional Mongolian history and should not reflect contemporary traditional Mongolian orthography only. For example; the feminine variants of traditional Mongolian syllable “QA+E” have at least 3 forms (ᠬᠢᠬᠢᠨ), but this scheme mentions none; the final forms of traditional Mongolian letter NA have at least 6 forms (ᠨᠠᠨᠠ, ᠨᠠᠨᠢ, ᠨᠠᠨᠣ, ᠨᠠᠨᠤ, ᠨᠠᠨᠦ, ᠨᠠᠨᠦ), but this scheme has 2 forms (ᠨᠠᠨᠠ) only. If Uighur Mongolian is included, the variants of traditional Mongolian will become even complex.

Third, this scheme disrupts the traditional Mongolian alphabet systems and handles pairs of individual letters as “several variants of one letter”. The traditional Mongolian alphabet is an indispensable element of the orthography. For example; letters ᠠ, ᠡ, ᠢ and ᠣ are individual letters and no need to be differentiated by FVSs. This scheme specifies that ᠠ and ᠡ” do not need FVSs while ᠢ and ᠣ need FVSs in any conditions, obviously not conformed to the Mongolian alphabet system and rules. Both Chinese national standard of *GB/T 26226-2010 Information Technology - Mongolian Presentation Forms Character Set and Use Rules of Controlling Characters* and the Mongolian national standard *MNS 4932: 2000 монголжисин бичгийн кодыг хэрэглэх дүрэм (Use of Mongolian Character Encoding)* specify that the FVSs of Mongolian scripts is used for differentiating the different free variants of one letter under the same conditions. Unfortunately, the specifications in current scheme not only disrupt the traditional Mongolian alphabet, while the use of the frequency of FVS is dramatically increased! According to statistics and on the basis of specifications in this scheme, inputting the 19,400 words of essay needs to be input 2,836 more FVSs just for ᠠ ᠣ letters!

Noted

It is not totally clear what is objected. The Mongolian encoding scheme was not a new part of this edition. Mongolian has been part of ISO/IEC 10646 for a long time and has been unchanged for many editions of the Standard. What was new in this Committee Draft was a modified of variation sequences based on a rough consensus by a large group of Mongolian experts from various constituencies as well as the introduction of the presentation forms (isolate, initial, medial, final) in the code chart to help the understanding of these variation sequences.

Many of the points raised here go much beyond of what can be described in a standard like ISO/IEC 10646 and should be described in a technical note, or added to the Unicode Standard which contains much more technical details about implementation. In all cases, should new text be added to 10646, it needs to be proposed by China and Mongolia to be actionable.

A Committee Draft is an opportunity to discuss new ideas and allow various experts to voice feedback on proposed changes.

Based on this feedback, it seems prudent to reverse the code charts to the previous version of the variation sequences as originally created, while still preserving the default presentation forms. This gets back to the original proposal as it was adopted in 10646 and never objected since.

There is however a large consensus that the current state of affair is not endorsed by all communities and all parties are encouraged to develop a consensus. Until that is done, the code charts will stay as before (except for the addition of the presentation form).

T2. Technical errors

(1) Lack of Unifications of the variants among 4 sub-scripts

The unifications of the variants of 4 sub-scripts (Traditional Mongolian, Todo, Sibe and Manchu) are complicated and important. For instance, it is common that 1 variant is used by 1 sub-script or shared by 2, 3, or 4 sub-scripts. Without the unifications of 4 sub-scripts, the system implementation and higher level process of language information such as TTS, OCR, MT, etc. are impossible.

Noted

It is not clear what the objection is. The 4 'sub-scripts' or more precisely writing systems are unified in the Mongolian block and have been also unified in all know fonts. This allows font rendering to be writing system aware and render differently.

If there are issue with the current standard (please refer to the 4th edition of ISO/IEC 10646), a clear and explicit proposal for desirable changes should be made.

(2) Lack of Mandatory Ligatures

The mandatory ligatures of Mongolian scripts are not covered in the scheme. The mandatory ligatures of Mongolian scripts are important and indispensable in the system implementation of Mongolian character encoding and in font design.

Noted

This is probably beyond the scope of the standard, see disposition of comment T1.

If there are issue with the current standard (please refer to the 4th edition of ISO/IEC 10646), a clear and explicit proposal for desirable changes should be made.

(3) Lack of Non-mandatory Ligatures

The non-mandatory ligatures of Mongolian scripts are not covered in the scheme. The non-mandatory ligatures are important and indispensable in the system implementation of Mongolian character encoding and in font design.

Noted

This is probably beyond the scope of the standard, see disposition of comment T1.

If there are issue with the current standard (please refer to the 4th edition of ISO/IEC 10646), a clear and explicit proposal for desirable changes should be made.

(4) Lots of Mistakes in Rules of FVSs

Using FVSs (Free Variation Selectors) or not and which FVS should be selected depend on the theories, rules and practices of the Mongolian orthography. The scheme does not obey the above principles and leads many errors and omissions in the usage rules of the FVS within a word.

Noted

This is probably beyond the scope of the standard, see disposition of comment T1.

If there are issue with the current standard (please refer to the 4th edition of ISO/IEC 10646), a clear and explicit proposal for desirable changes should be made.

(5) Lack of Character Sequences of Single Presentation Forms

This revision does not contain the whole character sequences of single presentation forms. If the complete character sequences of single presentation forms are not given, it will influence the input, storage, display and transfer of the single presentation forms.

The above mentioned five contents can be found in Chines national standard *GB/T 26226-2010* and Mongolian national standard *MNS 4932: 2000*.

Not accepted

There is no need to describe how to describe character sequences of single presentation forms. The name of each presentation form (isolate, initial, medial, and final) is self-explanatory and determines which form is presented when present in a word on the appropriate position. When these presentation forms are shown in isolation, it is possible to use U+200D ZERO WIDTH JOINER before and/or after; but this is just a mechanism to override the default behavior which would otherwise show the character in isolated form.

(6) Incomplete and Unpractical term -“Context Driven”

Fifties of rather unclear terms-“Context Driven” and “needed to override default context” are mentioned in the scheme and are not given where people can get them.

Accepted

These annotations will be removed, unless new information is provided

(7) Issues on Ali Gali Letters

Ali Gali letters’ presentation forms of three sub-scripts (traditional Mongolian, Todo and Manchu) of the scheme are incomplete.

Noted

Please provide the information so that these presentation forms could be improved.

(8) Issues on Cyrillic

The comparison of traditional Mongolian letters and Cyrillic small letters of the scheme makes no sense and is incomplete (1834 → 11/4) and incorrect (1824 → Y[u]).

Noted

It seems difficult to accept that comparison of traditional Mongolian letters and Cyrillic small letters make no sense when these two scripts are used to write Mongolian.

If the entry for 1834 and 1824 are either incomplete or incorrect, appropriate suggestion for the correction and completeness would be appreciated.

T3. GB/T 26226 and MN 4932 shall be respected

The content of GB/T 26226 and MNS 4932 was jointly developed by both China and Mongolia as guidelines for UCS and are being widely supported, thus the both countries should have been consulted for any revision of Mongolian encoding.

The China national standard GB 26226-2010 is being widely supported for fonts, inputting methods, offices, OS and publishing systems by Microsoft, Founder, Jade Bird Huaguang, China Standard Software, Menksoft, ISCAS, Inner Mongolia University and other developers, tremendous data of corpuses (about 23 million words), Knowledge bases (about 17.9 million records), E-books (1.79 million entry), web sites (145 thousand documents), governmental OAs are accumulated. This revision will introduce incompatibility thus cannot be accepted by developers and users. If the scheme has become the standard, it will lead incompatibility of mass data and bring troubles, losses and disasters.

Since the MNS 4932 was released in 2000, tremendous data resources generated by the products based on the standard have been accumulated, such as the Online Great Dictionary of Mongolian Language (ordered by the President Office), governmental documents, web sites, e-books and others.

Noted and partially accepted

There was no intent to ‘revise’ the Mongolian encoding. The intent of the new charts was just to reflect the use of variations sequences as mutually agreed by a large group of experts. As noted in the disposition of Comment 1, the text of the standard will be reversed back to the previous state.

T4. Suggestions on the structure of Mongolian character encoding

a) The “Standardized Variation Sequences” of Mongolian in UCS 2014 is adopted in principle. The sequences of the presentation forms need to be specified correctly. For instant, the second initial form “1820 180B” of Mongolian letter A (1820) should be revised as “200D 1820 180B 200D” and the second final form “1820 180B” should be revised as “200D 1820 180B”.

Not accepted

As noted in the answer to comment T2.5) this is not necessary. These U+200D code points are only needed when these presentation forms are shown in isolation and using them is indeed the correct solution in that context.

However, they are not needed when the sequences are located inside a word.

b) The usage rules of FVS within a word of different sub-scripts need to be written respectively because it is complicated, unclear and long-winded to have all rules together for all sub-scripts. The draft and technical document of the 4 Mongolian scripts (traditional Mongolian, Todo, Manchu and Uighur Mongolian scripts) will be submitted in the name of China and Mongolia later.

Noted

The contribution will be welcome and acted on when received.

c) It will be many new contents in the joint proposal on the 4 different Mongolian scripts. There are 20 presentation forms and particular writing forms from 1589 to 1949. The orthography of classical Todo is added in the Todo script part (including more than 50 presentation forms and particular writing forms). The peculiar contents of Manchu orthography (including more than 30 presentation forms and particular writing forms) and the diverse contents of Uighur Mongolian (including more than 30 presentation forms and particular writing forms) are added.

Noted

Technical comments on New Tai Lue

The character names of 19AA and 19AB should be changed to NEW TAILUE LETTER HIGH SVA and NEW TAILUE LETTER LOW SVA respectively.

Explanation: The two characters are both labialized consonant clusters same as the characters 19A6, 19A7, 19A8 and 19A9. It means that the lips are rounded when the consonants are articulated, making the consonants have the feature of round vowels. Therefore the naming should be coherent. The name SUA tends to be confused with the combination of consonant and round vowel U, as well as the combination of U and other vowels. When the consonant is combined with the round vowel U, the consonant should be pronounced first and then the U. While the function of the labialized consonant is to pronounce the consonant with lips rounded.

Partially accepted

It is not possible to change character names once encoded in the Standard. Two mechanisms are possible to achieve some of the intended result:

a) *Introduce a character name alias preceded by ‘※’ in the name list for both characters as in:*

19AA ɔꞀ NEW TAI LUE LETTER HIGH SUA
※ NEW TAI LUE LETTER HIGH SVA

19AB ɔꞁ NEW TAI LUE LETTER LOW SUA
※ NEW TAI LUE LETTER LOW SVA

b) *Introduce an annotation in the name list describing the issue such as:*

19AA ɔꞀ NEW TAI LUE LETTER HIGH SUA
• a better name for this character and the following would be sva because they are both labialized consonant clusters as the previous characters

19AB ɔꞁ NEW TAI LUE LETTER HIGH SUA

Technical comment on Miao

The following glyphs should be corrected as below:

16F2C ɛ MIAO LETTER NYA

16F2D ɛ MIAO LETTER NYHA

16F2E Ƀ MIAO LETTER TSHA

16F2F Ƀ MIAO LETTER DZHA

16F30 Ƀ MIAO LETTER YI TSHA

16F32 Ƀ MIAO LETTER REFORMED TSHA

16F31 Ƀ MIAO LETTER YI DZHA (Glyph position correct; no need to change)

Explanations:

- 16F2C  : This character is used in the *Normalised* but not the *Traditional Miao Orthography*. Moreover, the glyph shape per Miao scholars' design is .
- 16F2D  : This character is used in neither the *Traditional* nor the *Normalised Miao Orthography*. According to church pastors and Miao scholars, it is used amongst the Yi group (the White Yi of Xundian, Kunming, Yunnan Province, China), but no pronunciation information is available. From the spelling rules of the [Traditional] Miao orthography, it is probably the voiced counterpart of 16F2C .
- 16F2E  : This character is used in both the *Traditional* and the *Normalised Miao Orthography*. The difference is that it represents two phones in the *Traditional Miao Orthography* but only one in the *Normalised Miao Orthography*. No information about pronunciation and usage amongst the Yi is available.
- 16F2F  : This character is used only in the *Traditional Miao Orthography*. It is the voiced counterpart of 16F2E .
- 16F30  : This character is used in neither the *Traditional* nor the *Normalised Miao Orthography*. No information about pronunciation and usage amongst the Yi is available.
- 16F32  : It is known that this character was once used amongst the Miao of Stone Gateway, Guizhou before 1949. It had the same pronunciation as *j* in Chinese pinyin and one of the phones of in the *Traditional Miao Orthography*.

Accordingly, based on verification of character origin and pronunciation, we agree to revise the positions of the six glyphs as shown below, but we would like more accurate pronunciation verification for those characters used amongst the Yi.

Accepted

See also comment T2 from US.

The changes are as follows from previous version:

16F2C → 16F32

16F2D → 16F2C

16F2E → 16F2D

16F2F → 16F2E

16F30 → 16F2F

16F31 unchanged

16F32 → 16F30

The request about pronunciation verification is noted.

Technical comment on CJK Unified Ideographs Extension F

(The detailed feedback is in the HKSAR review to IRG 2130 (which was a document containing the CJK Ext F chart): http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg45/IRGN2130_HKSAR_Review.pdf , some of the material is duplicated here. All these should be discussed by IRG; however the project editor has consulted with CJK experts and come with proposed dispositions. Various IRG documents mentioned in the following section can be found at <http://appsrv.cse.cuhk.edu.hk/~irg/index.htm> .)

1. Radicals and Stroke Counts

a) 2CF16, SC 3 (according to IRG 2105, Appendix 2)



Propose accepted

5.1 -> 5.3

b) 2CF25, SC 8 (according to IRG 2105, Appendix 2)

2CF25 乙 5.6 
JMJ-057893

Propose accepted

5.6 -> 5.8

c) 2CF36, SC 13, reference 55F6, 84FD, and 4EBF

2CF36 乙 5.12 
USAT-00615

Propose accepted in principle

5.12 -> 5.13, but 102.9 would be better (102: 田)

d) 2D02F, SC 10, reference #22 of IRG N954AR & IRG N1105

2D02F 儿 10.11 
GZ-0062201

Propose not accepted

It would be 10.11 -> 10.10. I was given the example of IRG N954AR #78 to keep current residual stroke count

e) 2D108, SC 12, reference #35 of IRG N954AR & IRG N1105

2D108 刀 18.11 
USAT-01853

Propose accepted

18.11 -> 18.12

f) 2D170, SC 4

2D170 卜 25.5 
JMJ-058577

Propose accepted

25.5 -> 25.4

g) 2D2E0, SC 15, reference 6A06, 7483, and 96E2

2D2E0 口 30.14 
USAT-03237

Propose accepted

30.14 -> 30.15

h) 2D332, SC 19

2D332 口 30.18 
USAT-00433

Propose accepted

30.18 -> 30.19

i) 2D4F2, SC 6, the SC of the residual component is 6 in KangXi dictionary, reference 2DD23, 2DA99

2D4F2 彳 40.7 
USAT-02978

Propose accepted

40.7 -> 40.6

j) 2D68D, SC 18, reference 50B2, 71AC, and 8D05

2D68D
广 53.17 廢
KC-06230

Propose accepted

53.17 -> 53.18

k) 2D8A8, SC 12, reference 2E282 and 2E8B6

2D8A8
手 64.13 備
USAT-05373

Propose not accepted

It would be 64.13 -> 64.12. I was given the example of IRG N954AR #36 to keep current residual stroke count

l) 2D8FC, SC 10

2D8FC
支 66.9 鼓
USAT-05561

Propose accepted

66.9 -> 66.10

m) 2D9A9, SC 9, see Appendix 1 of IRG N2105

2D9A9
日 72.8 曾
USAT-03442

Propose accepted

72.8 -> 72.9

n) 2DA1E, Radical = Sun 日 (R72), reference U+66A2 暢 and U+7545 暢, [and 2D9C2]

2DA1E
日 73.7 𠄎
USAT-04132

Propose accepted

73.7 -> 72.7 (radical changed, no change in SC)

o) 2DA45, SC 9

2DA45
月 74.10 滕
USAT-04120

Propose not accepted

It would be 74.10 -> 74.9. I was given the example of IRG N954AR #35 to keep current residual stroke count

p) 2DA5F, SC 4, reference 4F3C, 59D2, and 62DF

2DA5F
木 75.3 杝
JM.J-059709

Propose accepted

75.3 -> 75.4

q) 2DB83, SC 23, on two nearly identical components in lower right, the one on the left has one less stroke

2DB83
止 77.24 𠄎兔
KC-05216

Propose accepted

77.24 -> 77.23

r) 2DB88, SC 2, SC of left component (radical) is 5 in KangXi dictionary, reference 2DB85, 400E, and 5DE7

2DB88
歹 78.3
𣦵
USAT-03735

Propose accepted

78.3 -> 78.2

s) 2DCF5, SC 16, SC of lower right component is 6 in KangXi dictionary, reference 2DD23

2DCF5
水 85.17
漕
USAT-04005

Propose accepted

85.17 -> 85.16

t) 2DDE5, SC 6, SC of right component is 6 in KangXi dictionary, reference 5905, 9004, and 964D

2DDE5
牛 93.7
犇
USAT-03103

Propose accepted

93.7 -> 93.6

u) 2D49E, SC 13

2D49E
女 38.12
姪
USAT-01951

IRG discussion

This would be 38.12 -> 38.13. That changes was not in HKSAR feedback and is out of sequence, it could be a typo for 2DE9E which is part of the HKSAR feedback and is not listed here.

v) 2DF35, SC 12

2DF35
疒 104.14
瘡
KC-05441

IRG discussion

This would be 104.14 -> 104.12. Another opinion is 104.13.

w) 2DFDE, SC 12

2DFDE
目 109.11
瞞
KC-06555

Propose accepted

109.11 -> 109.12

x) 2E0E1, SC 11, reference 5368, 21A9D, and 25801

2E0E1
禾 115.10
稿
KC-05509

Propose accepted

115.10 -> 115.11

y) 2E0F9, SC 16, reference 417B, 6A06, and 7483

2E0F9
禾 115.15
𪔑
USAT-80094

Propose accepted

115.15 -> 115.16

z) 2E113, SC 11, reference 5368, 21A9D

2E113 窩
穴 116.10
GCY-2961.00

Propose accepted

116.10 -> 116.11

aa) 2E172, SC 8, reference of lower right component in KangXi dictionary is 5, reference 5B64, 72D0, and 26C44

2E172 籐
竹 118.9
USAT-04888

Propose accepted

118.9 -> 119.8

ab) 2E207, SC 13, see appendix 1 of IRG N2105

2E207 允米
米 119.14
USAT-02077

Propose accepted

119.14 -> 119.13

ac) 2E28B, SC 4, reference 7F52

2E28B 今
网 122.5
KC-05582

Propose accepted

122.5 -> 122.4

ad) 2E2B8, SC 6, [probably a typo for 2E2B2 (2E2B8 RSC is 123.9), 2E2B2 is part of the HKSAR feedback, 2E2B8 is not]

2E2B2 羴
羊 123.7
GPGLG-3035

Propose accepted in principle

123.7 -> 123.6

ae) 2E37A, SC 10 [another typo, should say SC 9], SC should be 9 if outer component is taken as the radical

2E37A 興
白 134.10
JMJ-060094

Propose accepted in principle

134.10 -> 134.9

af) 2E38A, SC 4, IRG has resolved in appendix 2 of IRG N 2105 that the upper component SC is 5, but it is 4 in KangXi, and in reference 65E2, 65E3, 2312D, 2312E, and 2312F, the minor stroke is not counted.

2E38A 死
殳 136.5
JMJ-058426

Propose accepted

136.5 -> 136.4

ag) 2E3A7, SC 5, the SC of the lower component is 5 in KangXi

2E3A7 𪗇
冊 140.4
USAT-03713

Propose accepted

140.4 -> 140.5

ah) 2E4F1, SC 11

2E4F1 𪗇
虍 141.10
KC-06741

Propose accepted

141.10 -> 141.11

ai) 2E4F4, SC 9, same rationale as for 2E38A

2E4F4 𪗇
虍 141.10
USAT-80110

Propose accepted

141.10 -> 141.9

aj) 2E685, SC 4, same rationale as for 2E38A

2E685 既
貝 154.5
KC-03956

Propose accepted

154.5 -> 154.4

ak) 2E713, SC 11, reference 6A06, 7483, and 96E2

2E713 躄
足 157.10
USAT-01708

Propose accepted

157.10 -> 157.11

al) 2E882, SC 3 [*probably a typo for 2E822 (2E882 RSC is 167.6), 2E822 is part of the HKSAR feedback, 2E882 is not, also out of sequence*], reference 2E820

2E822 𪗇
酉 164.2
JM.J-058883

Propose accepted in principle

164.2 -> 164.3

am) 2E84B, SC 12

2E84B 𪗇
酉 164.11
KC-06837

Propose accepted

164.11 -> 164.12

an) 2E8E4, SC 12

2E8E4 鍍
金 167.14
KC-05832

IRG discussion

This would be 167.14 -> 167.12. Another opinion is 167.13.

ao) 2E914, SC 4, same rationale as for 2E38A

2E914 閼
門 169.5
USAT-03837

Propose accepted

169.5 -> 169.4

ap) 2E917, SC 4, the SC of the enclosed component is 5 in KangXi, reference 244F0

2E917 閼
門 169.5
USAT-01965

Propose not accepted

It would be 169.5 -> 169.4, but the KangXi SC value is not supporting the change and 244F0 is different

aq) 2EA47, SC 5

2EA47 飼
食 184.6
KC-05898

Propose accepted

184.6 -> 184.5

ar) 2EA85, SC 4, same rationale as for 2E38A

2EA85 駣
馬 187.5
KC-07176

Propose accepted

187.5 -> 187.4

as) 2EB4B, SC 9 or 10? ref. U+2DAD4

2EB4B 𪚗
鳥 196.8
USAT-06651

IRG discussion

Propose no change: 𪚗 (2) + 一 (1) + 出 (5)

at) 2EBB3, SC 5 or 6? Ref. U+214A1, U+2503B, U+2503D, U+2503E, U+2A1DE, U+ 2A50F (only this one supports current counting method.)

2EBB3 𪚗
鼓 207.5
KC-06965

IRG discussion

Propose 207.5 -> 207.6 (majority rule)

au) 2EBB6, SC 12? Ref. U+214A1, U+2503B, U+2503D, U+2503E, U+2A1DE, U+ 2A50F (only this one supports current counting method.)

2EBB6 𪚗
鼓 207.11
KC-07188

IRG discussion

Propose 207.11 -> 207.12 (majority rule)

In addition to the code points mentioned above, the HKSAR feedback about IRG N2130 contains feedback for the following code points:

av) 2CF19, SC=3?

2CF19 地
乙 5.4
JM.J-056867

IRG discussion

Propose to stay as it is (5.4), unlike 2CF16 the oblique stroke does not connect with the vertical stroke.

aw) 2D2B3, SC =13

2D2B3 啤
口 30.12
USAT-05759

IRG discussion

Propose to do 30.12 -> 30.13 as suggested.

ax) 2DA11, SC=16?, reference 2A813, 2B249, 4B1D, and 4D43

2DA11 蹟
日 72.17
GZ-2501101

IRG discussion

Propose to do 72.17 -> 72.16 as suggested.

ay) 2DE95, SC =17

2DE95 璽
玉 96.16
USAT-05286

IRG discussion

Propose to do 96.16 -> 96.17 as suggested (or the two vertical strokes should be merged into one).

az) 2DE9E, SC 13, reference 6A9B, 6FC4, 203C0 (may be part of the Chinese comment if 2D49E is a typo)

2DE9E 𪗗
瓜 97.12
GZYS-00321

IRG discussion

Propose to do 97.12 -> 97.13 as suggested.

ba) 2E0CD, SC 9, reference 5835, 7779, and 8AF8

2E0CD 稽
禾 115.8
KC-06592

IRG discussion

Propose to do 115.8 -> 115.9 as suggested.

bb) 2E495, SC=15, the SC of the middle component is 12 in KangXi, reference 655D, 5E63, and 853D

2E495 藜 藜
艸 140.14
KC-06735 USAT-02594

IRG discussion

Propose to do 140.14 -> 140.15 as suggested.

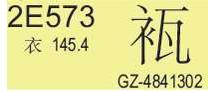
bc) 2E4A3, SC=16

2E4A3 藓
艸 140.15
USAT-06293

IRG discussion

Propose to do 140.15 -> 140.16 as suggested.

bd) 2E573, SC=5, the SC of the right component is 5 in KangXi, see also appendix 1 of IRG N2105, reference 2E5F3



IRG discussion

Propose to do 145.4 -> 145.5 as suggested.

be) 2E5CD, SC=9



IRG discussion

Propose to do 146.8 -> 146.9 as suggested.

bf) 2E7AE, SC=9



IRG discussion

Propose to do 162.8 -> 162.9 as suggested.

bg) 2E9A7, SC=10



IRG discussion

Propose to do 173.9 -> 173.10 as suggested.

2. Wrong Positions

KC-06578, 衤 令 should be moved from U+2D1D1 to U+2E068 . Its r/s value should be 113.5. See IRGN2125 Consolidated Review.

KC-07044 木名 should be moved from U+2D394 to U+2DA7C . Its r/s value should be 75.6. See IRGN2125 Consolidated Review.

Accepted

See also comments T2 from Japan, T7 and T8 from UK, TE5 and TE6 from US.

3. Fonts to be Improved

For references of above comments on CJK_F, see the HKSAR's Feedback on IRGN2130 from IRG web site.

2CFDF

2D0CF

2D0D0

2D102

2D16F

2D171

2D5F8

2D73E [probably a typo for 2D7E3]

2DA69

2DA90

2DE2F

2DE95

2DF02
2DFC5 (USAT font to be improved)
2E41B
2E608

IRG discussion

Compared to the HKSAR feedback, two entries are missing: 2DAD4, 2E72C; and one entry is in the wrong category (2DE95 is a SC issue and is discussed above).

It looks like 2DF02 is probably intentional and should not be modified.

2DF02 疋
疋 103.4 疋
JMJ-058113

2DFC5 USAT entry needs to be fixed to be similar to the K entry:

2DFC5 睽 睽
目 109.8 睽 睽
KC-05456 USAT-04573

:

India: Negative

Technical comments

IN.1. Page 1358 to 1360, Tamil Supplement 11FC0-11FFF.

Symbols proposed for the Tamil Supplement block in the code points between 11FC0-11FFF have major technical errors.

Tamilnadu Government / India has proposed amendments to these symbols. The amendment proposal is under the review of UTC and published for the Public review as well.

<http://www.unicode.org/L2/L2016/16062-tamil-frac-sym-fdbk.pdf> .

Proposed change by India

Not accepted for encoding now.

Withdraw the original proposal that proposes the symbols in the code points between 11FC0-11FFF.

The code points between 11FC0-11FFF should wait and accommodate the correct symbols proposed through the Amendment from Tamilnadu Government / India after the review of UTC.

Accepted

*The proposal for encoding the content of the Tamil Supplement block is postponed, therefore removed from this edition and will **not** be moved to the next amendment for ballot.*

IN.2. Page 145, Tamil 0B99

Annotation of 0B99 has reference to 11FD5 which is proposed to be deferred from encoding.

See "Proposed change" of Comment "IN 1".

Proposed change by India.

The annotation "→ 11FD5 ஈஃ tamil sign muuvulakku" has to be removed.

Accepted

IN.3. Page 145, Tamil 0BA4

Annotation of 0BA4 has reference to 11FD7 which is proposed to be deferred from encoding.

See "Proposed change" of Comment "IN 1".

Proposed change by India.

The annotation "→ 11FD7 ஈஃ tamil sign mukkuruni" has to be removed.

Accepted

IN.4. Page 145, Tamil 0BB3

Annotation of 0BB3 has reference to 11FD7 which is proposed to be deferred from encoding.

See "Proposed change" of Comment "IN 1".

Proposed change by India.

The annotation "→ 11FD7 ஈஃ tamil sign mukkuruni" has to be removed.

Accepted

IN.5. Page 146, Tamil Numerics

Notes under the paragraph/title "Tamil Numerics" has references to the Tamil Supplement block which is proposed to be deferred from encoding.

See "Proposed change" of Comment "IN 1".

Proposed change by India

The notes " Tamil fractions are encoded in the Tamil Supplement block at 11FC0-11FFF." has to be removed.

Accepted

IN.6. Page 146, Tamil

Notes under the paragraph/title "Tamil symbol" has references to the Tamil Supplement block which is proposed to be deferred from encoding.

See "Proposed change" of Comment "IN 1".

Proposed change by India

The notes "More symbols are encoded in the Tamil Supplement block at 11FC0-11FFF." has to be removed.

Proposed accepted

IN.7. Page 145, Tamil 0BFA

Annotation of 0BB3 has reference to 11FF1 which is proposed to be deferred from encoding.

See "Proposed change" of Comment "IN 1".

Proposed change by India.

The annotation "→ 11FF1 ழ. tamil traditional number sign" has to be removed.

Accepted

Japan: Negative

Technical comments

If T.1 and T.2 are accommodated, Japan NB changes its vote to Yes.

T1. Page 2489-2582, Clause 33 – “CJK Unified Ideographs Extension F”

At the last WG2 meeting in Japan, WG2 issued the following recommendation. (Recommendation M64.11)

"WG2 recommends that IRG reviews its CJK unification rules to minimize the number of glyph variants that are coded as separate characters."

Following this recommendation, IRG reviewed CJK Extension F at the last IRG meeting in Hong Kong and concluded to unify some CJK characters with the glyph variants that are separately encoded.

However, after IRG meeting, Japan NB found more CJK F characters to be unified with the glyph variants based on the same principle.

See the following table showing the CJK glyphs to be unified with. (No special meaning in the area highlighted in yellow.).

Proposed CJK F char	To be unified with		
1 2D0D2 刂 18.5 USAT-00085	206B9 刂 18.4 UCS2003 GKX-0137.08 T5-222E	2DAE0 木 75.10 USAT-02856	23624 木 75.11 UCS2003 GKX-0549.25 T4-4A3F
2 2D0D6 刀 18.6 GZ-4861102	52BD 刀 19.6 GE-2363 T4-2550 J1-3364 K2-2532	2DAFD 木 75.11 USAT-01188	6A7B 木 75.12 GK-8899 K2-3D3A
3 2D0E9 刀 18.8 GZ-4082201	2B9AF 刀 19.8 GZFY-00521	2DB0F 木 75.12 KC-07047	69FB 木 75.11 G3-416A HB2-E5E2 T2-512B J0-4450 K0-5027
4 2D10E 刀 18.11 USAT-05989	2D13D 刀 19.11 USAT-01966	2DB54 欠 76.8 JMJ-057882	238C7 欠 76.7 UCS2003 GHZ-80027.22
5 2D136 刀 19.9 USAT-60035	20778 刀 18.9 UCS2003 GHZ-60010.24	2DB8C 歹 78.4 USAT-05410	239CA 歹 78.4 UCS2003 GKX-0579.17 T5-2660
6 2D2E5 口 30.14 USAT-05290	210A0 口 30.15 UCS2003 GKX-0213.07 T4-5A79	2DB90 歹 78.5 JMJ-057905	6B81 歹 78.4 G0-6962 H-8DD1 T3-2958 J4-2F7E
7 2D421 爻 35.5 USAT-02739	225BB 心 61.3 UCS2003 GKX-0377.07 T6-275C	2DBB8 爻 79.7 JMJ-057919	3C7F 爻 79.6 G5-5063 T3-316B
8 2D44B 大 37.6 KC-00818	77E9 大 111.5 G0-3E58 HB1-AF78 T1-567B J0-368B K0-4F3B V1-605D	2DBC0 歹 79.10 JMJ-057926	6BBF 歹 79.9 G0-365E HB1-B765 T1-6A5B J0-4542 K0-6E7C V1-5848
9 2D584 山 46.3 USAT-00821	21D45 山 46.3 UCS2003 J4-2838	2DBC2 歹 79.11 JMJ-057928	6BC4 歹 79.10 GE-2F54 HB2-E26A T2-4832 J1-4635 K1-5830
10 2D6E1 彳 60.5 JMJ-057485	5F79 彳 60.4 G0-5258 HB1-A7D0 T1-4A32 J0-4C72 K0-6635 V1-545C	2DBF6 氏 83.4 JMJ-057936	6C13 氏 83.4 G0-4325 HB1-AA5D T1-4E40 J0-5D62 K0-686C V1-5858
11 2D818 手 64.5 JMJ-057554 USAT-03285	6295 手 64.4 G0-4D36 HB1-A7E8 T1-4A4D J0-456A K0-7761 V1-565A	2DEF3 田 102.12 JMJ-058106	2DEF8 田 102.14 JMJ-058108
12 2D81D 手 64.5 USAT-05471	22AE4 手 64.6 UCS2003 GKX-0426.07 T6-3152	2E0DD 禾 115.9 USAT-02356	2581E 禾 115.9 UCS2003 GHZ-42617.16
13 2D8B4 手 64.14 USAT-01809	6503 手 64.15 G3-4B4F HB2-EFD6 T2-6162 K2-3746	2E100 穴 116.5 USAT-02971	4193 穴 116.4 GKX-0863.12 T3-2E46
14 2D99C 日 72.8 JMJ-059691	666B 日 72.8 G3-4771 H-93D3 T3-3B7D J14-2E26 K0-7678	2E168 竹 118.9 USAT-60250	25C4D 竹 118.10 UCS2003 T5-5658
15 2DA4C 月 74.12 JMJ-037911	2339F 月 74.12 UCS2003 T5-546D	2E1CF 米 119.4 USAT-03923	7C87 米 119.4 G6-3F5D T3-3330 J1-632D K2-5146
16 2DA99 木 75.7 JMJ-057740	6840 木 75.6 G0-686E HB1-AEE5 T1-5629 J0-5B5C K0-4B7A V1-6971	2E1D3 米 119.5 JMJ-058287	25E4F 米 119.5 UCS2003 GKX-0908.12 T6-4C63

33	2E232 緝 糸 120.9 USAT-01467	27322 緃 緃 緃 虫 142.8 UCS2003 GHZ-1088.22 T5-4961
34	2E295 罔 网 122.9 JMJ-058343	262DF 罔 罔 罔 网 122.9 UCS2003 GHZ-0948.36 T7-2675
35	2E304 聲 耳 128.12 JMJ-058370	8072 聲 聲 聲 聲 聲 聲 耳 128.11 G1-4979 HB1-C16E T1-7478 J0-6661 K0-6122 V1-6363
36	2E314 𦞑 𦞑 肉 130.4 JMJ-058379 USAT-80289	809E 𦞑 𦞑 𦞑 肉 130.3 GE-3962 T4-2470 J1-556E
37	2E330 羸 肉 130.10 JMJ-058394	267C0 羸 羸 羸 肉 130.10 UCS2003 GHZ-32104.04 T6-6070
38	2E341 膈 肉 130.12 JMJ-058398	8188 膈 膈 膈 膈 膈 肉 130.10 G0-6B75 HB1-8B4A T1-6A52 J0-672D K0-4C2E
39	2E368 臺 至 133.4 JMJ-060086	2693C 臺 臺 臺 至 133.5 UCS2003 GHZ-42816.01 T4-3456
40	2E40B 蔡 艸 140.10 JMJ-060130	8521 蔡 蔡 蔡 蔡 蔡 艸 140.11 G0-324C HB1-8D62 T1-8E58 J0-6671 K0-7379
41	2E484 裴 衣 140.13 USAT-03743	2E5AB 裴 衣 145.10 KC-06763
42	2E5E8 颯 見 147.14 JMJ-058726	2789B 颯 颯 見 147.14 UCS2003 GHZ-63676.07
43	2E615 說 言 149.6 JMJ-058743	8AAA 說 說 說 說 言 149.7 GE-3D73 HB1-8BA1 T1-6B29 K0-6063
44	2E63A 諫 言 149.10 USAT-00800	27AD5 諫 諫 諫 言 149.11 UCS2003 GHZ-1177.12 T4-5D68
45	2E752 輟 車 159.5 JMJ-058827	8ED7 輟 輟 輟 車 159.4 GE-3F62 HB2-07DF T2-3962
46	2E769 輟 車 159.9 JMJ-060221	2833A 輟 輟 車 159.8 UCS2003 H-8ECA
47	2E82F 醜 酉 164.5 KC-06802	2882E 醜 醜 醜 酉 164.6 UCS2003 GHZ-63581.11 T6-694D
48	2E9E0 軌 革 177.4 USAT-03573	29273 軌 軌 革 177.4 UCS2003 GHZ-74327.20

Proposed change by Japan.

- Please delete the following CJK F characters, because either they are unified with other CJK unified character already encoded before CJK F or they are unified with other CJK F character having same source category (#4 and #27 of the table on left-hand side “Comments” column). Note the list is sorted in the order of code point.
U+2D0D2, U+2D0D6, U+2D0E9, U+2D10E, U+2D136, U+2D2E5, U+2D421, U+2D44B, U+2D584, U+2D6E1, U+2D818, U+2D81D, U+2D8B4, U+2D99C, U+2DA4C, U+2DA99, U+2DAE0, U+2DAFD, U+2DB0F, U+2DB54, U+2DB8C, U+2DB90, U+2DBB8, U+2DBC0, U+2DBC2, U+2DBF6, U+2DEF3, U+2E0DD, U+2E100, U+2E168, U+2E1CF, U+2E1D3, U+2E232, U+2E295, U+2E304, U+2E314, U+2E330, U+2E341, U+2E368, U+2E40B, U+2E5E8, U+2E615, U+2E63A, U+2E752, U+2E769, U+2E82F, U+2E9E0, U+2EB02
- Please unify U+2E484 (#41 of the table) with U+2E5AB, so that U+2E5AB has the two glyphs from KC and USAT.

IRG discussion

The request seems reasonable but should be endorsed by IRG.

T2. Clause 33 – “CJK Unified Ideographs Extension F”

Following glyphs are placed on same code point by mistake.

2D1D1 𠄎 禛
口 30.5 JMJ-059415 KC-06578

2D394 格 址
土 32.6 KC-07044 USAT-03073

Proposed change by Japan.

KC-06578 should be moved to U+2E068.

2E068 禛
示 113.5 JMJ-059958

KC-07044 should be moved to U+2DA7C.

2DA7C 格
木 75.6 GZJW-01933

Accepted

See also comments CJK.2 from China, T7 and T8 from UK, TE5 and TE6 from US.

This was a production error. The issue with KC-06578 and KC-07044 had to do that the editor was given incorrect indexes in the original Korean font.

T3. Page 2652, Sub-clause A.5.10 390 MOJI-JOHO-KIBAN IDEOGRAPHS-2016

As described in Editor's Note, the file "JMJKI-2016.txt" should be consistent with the character set of CJK Extension F at the publication of the 5th edition of this International Standard.

Japan NB separately sent "JMJKI-2016.txt" file to the project editor based on the current CJK Extension F just in case.

Proposed change by Japan.

Please maintain the contents of file "JMJKI-2016.txt" to be correspondent with the character set of CJK Extension F at the publication of the 5th edition of this International Standard

Proposed accepted in principle

Given that the code allocation for CJK Extension F will be changed by these dispositions, the file JMJKI-2016.txt will need to be updated again. However, this could be fixed during the ballot time for the 5th edition in order to be available before publication.

Editorial comments

E1. Several throughout the whole document

We see "CJK Unified ideograph" (capital "U") and "CJK unified ideograph." (small "u")

Proposed change by Japan

Please make the case of "CJK Unified ideograph" be consistent. As for the block name, it would be "CJK UNIFIED IDEOGRAPH".

Accepted

E2. Page 31, 32 Sub-clause 23.1 List of source references

[Ed. Typo]

Proposed change by Japan

"Note 1", "Note 2" and "Note 3" should be "NOTE 1", "NOTE 2" and "NOTE 3."

Accepted

E3. Page 32 Sub-clause 23.1 List of source references

[Ed. Typo]

Proposed change by Japan

J3A JIS X 0213:2004 level-3 addendum from JIS JIS X 0213:2000 level-3

should be

J3A JIS X 0213:2004 level-3 addendum from JIS X 0213:2000 level-3

J13A JIS X 0213:2004 level-3 addendum from JIS X 0213:2000 level-3 replacing J1 characters

should be

J13A JIS X 0213:2004 level-3 addendum from JIS X 0213:2000 level-3 replacing J1 characters

Accepted

E4. Page 32 Sub-clause 23.1 List of source references

[Ed. Typo]

Proposed change by Japan.

“(see A.4.3 and A4.4)” in NOTE 3

should be

“(see A.4.3 and A.4.4)”

Accepted

E5. Page 2489 – 2582 Clause 33 “CJK Unified Ideographs Extension F”

Some USAT glyphs on CJK F are not correctly rendered. As shown below, there are many unfilled pixels at the position where the strokes are overlapped. We found this problem at some resolutions.



Proposed change by Japan.

USAT font should be fixed

Accepted in principle

Based on receiving such a font from SAT.

E6. Page 2637 Sub-clause Annex A.1 Collections of coded graphic characters

[Ed. Typo]

Proposed change by Japan.

1075 SUPPLEMENTAL ARROWS-C 1F800-1F8FF

should be

1075 SUPPLEMENTAL ARROWS-C 1F800-1F8FF

Accepted

Mongolia: Negative

The ballot comment from Mongolia is nearly identical to the technical comments concerning Mongolian made by China. Therefore, it is not repeated here. Please refer to the disposition of the Chinese comment T1 on Mongolian for the result on the CD content and all dispositions concerning Chinese comments on Mongolian T1 to T4.

Norway: Positive with comments

General comment:

G1. Sub-clause 4 Terms and definitions

ISO/IEC Directives Part 2 stipulates alphabetical ordering of terms and definitions as the “least preferred order”, while systematic order is the preferred order.

Proposed change by Norway.

Change order of terms and definitions

Accepted in principle

The project editor will take in consideration this input and after consultation with other experts in the Working Group will determine a preferred order for the list. The possible issue is cross reference from other standards.

G2. Sub-clause 4 Terms and definitions

ISO/IEC Directives Part 2 stipulates the use of the information category “Note # to entry” for use in the terminology section. The rules for that information category are slightly different from those of “NOTE”, in particular that normative information is permitted.

Proposed change by Norway.

Change “NOTE” to “Note # to entry” throughout

Accepted

For example, in the current term 4.1 base character, ‘NOTE 1’ becomes ‘Note 1 to entry’ and ‘NOTE 2’ becomes ‘Note 2 to entry’

Technical comment:

T1. Clause 32, Code Charts and lists of character names – Latin Extended-A

LATIN CAPITAL LETTER ENG

In the chart provided with the draft this character is correctly rendered in accordance with cultural expectations of the Sami languages (in particular Northern Sami). However the note text “glyph may also have appearance of large form of the small letter” is positively incorrect for Sami.

A large number of commercially available fonts have implemented the “note” in the character chart rather than the character form provided in the chart itself, making these fonts unacceptable for Sami. Since the character is listed under the heading “European Latin” it is assumed that the focus for these characters is on European languages and European usage.

We understand that a glyph with the form as indicated in the note is in use in other (non-European) languages, and that this glyph needs to be encoded. However, this needs to be done without causing encoding problems for languages that in fact are using and need the glyph Ŋ (and it was necessary to change font from Arial to Calibri to write the glyph correctly).

Standards Norway has been strongly requested by the Norwegian Ministry of Cultural Affairs to help ensure stable and appropriate encoding of text in our indigenous languages. Large quantities of text have been encoded using the character “014A” assuming that this is a unique representation of the CAPITAL LETTER ENG. It is noted that earlier versions of character set standards the character has been named “LATIN ... LETTER ENG (Sami)”, giving encoders the understanding that the correct form for Sami is indeed what is intended for code space 014A.

Proposed change by Norway.

Delete the note “glyph may also have appearance of large form of the small letter” and provide a separate space for “Latin capital letter ENG with shape as small letter ENG”..

Not accepted

The request amount to a dis-unification of the LETTER ENG. If Norway wants to entertain the proposal, it is up to Norway to create a proposal and follow the procedure for encoding of a new character. Until that is done and accepted, the informative annotation serves a very useful purpose.

United Kingdom: Negative

Technical comments start with ‘T’, and Editorial comments start with ‘E’:

E1. Page 12, Sub-clause 6.4, Naming of characters

“follows the rules given in 26.8 for Nushu characters, or”

For consistency with usage elsewhere in the standard, the bolded instance of “Nushu” should be written as “Nüshu”.

Proposed change by U.K.

Change “Nushu” to “Nüshu”.

Accepted

E2. Page 40, Sub-clause 25.2, Source reference file for Nüshu ideographs

“Source reference file for Nüshu ideographs”

Nüshu characters are not elsewhere referred to as “Nüshu ideographs”.

Proposed change by U.K.

Change to “Source reference file for Nüshu characters”.

Accepted

E3. Page 43, Sub-clause 26.8, Character names for Nushu characters

“Character names for Nushu characters

For Nushu characters the names are algorithmically constructed by appending their coded representation in their five hexadecimal digit notation to “NUSHU CHARACTER-”. For example, the first Nushu character has the name “NUSHU CHARACTER-1B100”.

For consistency with usage elsewhere in the standard, the bolded instances of “Nushu” should be written as “Nüshu”.

Proposed change by U.K.

Change “Nushu” to “Nüshu” in these instances.

Accepted

T4. Page 52, Clause 31 Structure of the Tertiary Ideographic Plane (TIP)

“The TIP (plane 03) is used for ancient ideographic scripts that are related to but not classified as CJK unified ideographs. No characters are currently encoded in the TIP.”

It is probable that the SIP will be filled and a new plane required to be assigned for CJK unified ideographs before any ancient ideographic scripts are ready for encoding. If this is the case, then it may be best to use the TIP for CJK unified ideograph extensions as well as or instead of ancient ideographic scripts. Therefore it is inadvisable to state categorically that the TIP is used for ancient ideographic scripts.

Proposed change by U.K.

Change to “The TIP (plane 03) is intended for CJK unified ideographs (unified East Asian ideographs) that are not encoded in the BMP or SIP. It may also include ancient ideographic scripts that are related to but not classified as CJK unified ideographs. No characters are currently encoded in the TIP.”

The note may be left unchanged.

Accepted

T5. Page 53, Clause 33, Code Charts and lists of character names – Nüshu

We sorted the Nushu repertoire according to the ordering rules noted in the code chart for Nushu, and the following 10 characters seem to be misplaced:

NUSHU CHARACTER-1B1E0 * tcie35 should be after NUSHU CHARACTER-1B1E1 * sie35

NUSHU CHARACTER-1B201 * sew35 should be after NUSHU CHARACTER-1B202 * lew44

NUSHU CHARACTER-1B22B * cyu44 should be after NUSHU CHARACTER-1B22C * tchyu21

NUSHU CHARACTER-1B237 * huow21 should be after NUSHU CHARACTER-1B238 * kuow44

NUSHU CHARACTER-1B239 * nuow42 should be before NUSHU CHARACTER-1B238 * kuow44

NUSHU CHARACTER-1B242 * liong42 should be before NUSHU CHARACTER-1B240 * tshiong35

NUSHU CHARACTER-1B253 * tshiu21 should be after NUSHU CHARACTER-1B255 * tci21

NUSHU CHARACTER-1B2C7 * lang33 should be after NUSHU CHARACTER-1B2C8 * tang13

NUSHU CHARACTER-1B2DD * tshew5 should be after NUSHU CHARACTER-1B2DE * lew33
NUSHU CHARACTER-1B2F2 * ku44 should be before NUSHU CHARACTER-1B2F0 * cyu35

Proposed change by U.K.

Reorder these ten characters

Accepted

T6. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2D0F0

U+2D0F0. The two source glyphs (JMJ-059378 and USAT-04376) have different left hand components, which would not seem to be unifiable.

Proposed change by U.K.

Encode JMJ-059378 and USAT-04376 as separate characters, JMJ-059378 under 8 strokes, and USAT-04376 under 9 strokes.

IRG discussion

The request seems reasonable but should be endorsed by IRG. For reference the chart entry looks like:

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

T7. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2D1D1

U+2D1D1. KC-06578 is misplaced, and should be unified with JMJ-059958 as U+2E068.

Proposed change by U.K.

Move KC-06578 to U+2E068 (with JMJ-059958).

Accepted

See also comments CJK.2 from China, T2 from Japan, TE5 and TE6 from US.

T8. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2D394

U+2D394. KC-07044 is misplaced, and should be unified with GZJW-01933 as U+2DA7C.

Proposed change by U.K.

Move KC-07044 to U+2DA7C (with GZJW-01933).

Accepted

See also comments CJK.2 from China, T2 from Japan, TE5 and TE6 from US.

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 牧 (See Zhonghua Zihai p. 1006).

Proposed change by U.K.

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

IRG discussion

The request seems reasonable but should be endorsed by IRG. For reference the chart entry looks like:

2DDC6
扌 90.4 收 牧
JMJ-057583 USAT-04734

T10. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2DF3B

U+2DF3B (USAT-04032) is under radical 104, but would be better placed under radical 61.

Proposed change by U.K.

Move U+2DF3B to radical 61, under 17 strokes.

IRG discussion

The request seems reasonable but should be endorsed by IRG. For reference the chart entry looks like:

2DF3B
彳 104.16 愆
USAT-04032

If accepted, the character would go after 2D7E2:

2D7E2 爆
心 61.17
USAT-03683

T11. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2DFC5
U+2DFC5. The source glyph for USAT-04573 has an incorrect radical (radical 72). The source for this character (http://21dzk.l.u-tokyo.ac.jp/SAT2012/T2157_.55.0886x11.html) shows that it should be radical 109.

Proposed change by U.K.

Correct the glyph for for USAT-04573 to have the correct radical (radical 109).

IRG discussion

See also comment T from China.

The request seems reasonable but should be endorsed by IRG.

T12. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2E048

U+2E048 (JMJ-058197) would be better placed under radical 86. Although this is a variant of U+7901 礁 (radical 112), the change in position of the four-dot fire element means that U+2E048 is best classified under radical 86.

Compare the analogous pair U+9EDE 點 (radical 203) and U+3E03 點 (radical 86).

Proposed change by U.K.

Move U+2E048 to radical 86, under 13 strokes.

IRG discussion

The request seems reasonable but should be endorsed by IRG. For reference the chart entry looks like:

2E048 礁
石 112.12
JMJ-058197

T13. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2E1F5

U+2E1F5. The two source glyphs (JMJ-058296 and USAT-00062) have different left hand components, which would not seem to be unifiable. Compare U+2DB74 (USAT-00811) and U+2E00D (USAT-02508) which have the same difference in left hand component, but are not unified.

Proposed change by U.K.

Encode JMJ-059378 and USAT-04376 as separate characters, JMJ-058296 under 9 strokes, and USAT-00062 under 11 strokes.

IRG discussion

The request seems reasonable but should be endorsed by IRG. For reference the chart entries look like:

2E1F5 米 119.9 JMJ-058296	𪛗 𪛗 USAT-00062	2DB74 止 77.11 𪛗 𪛗 USAT-00811	2E00D 矢 111.11 𪛗 𪛗 USAT-02508
--------------------------------	-------------------	---------------------------------------	--

T14. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2E321

U+2E321. The two source glyphs (JMJ-058387 and USAT-60296) have different left hand components, which would not seem to be unifiable.

Proposed change by U.K.

Encode JMJ-058387 and USAT-60296 as separate characters, both under 8 strokes.

IRG discussion

The request seems reasonable but should be endorsed by IRG. For reference the chart entry looks like:

2E321 胸 胸
肉 130.8
JMJ-058387 USAT-60296

E15. Page 2665, Annex G

Missing space between “and” and “Tangut”.

Proposed change by U.K.

Insert a space between “and” and “Tangut”.

Accepted

USA: Negative

(If TE.3, TE.5, and TE.6 are accommodated, the USNB will change its vote to yes.

Technical comments:

TE.1. Page 53, Clause 33, Code Charts and lists of character names – Sharada

Based on evidence in UTC document [L2/15-255](#), the glyphs for the following two characters need to be corrected:

111BA SHARADA VOWEL SIGN VOCALIC L

111BB SHARADA VOWEL SIGN VOCALIC LL

Proposed change by US:

Change the glyphs for U+111BA and U+111BB as follows:

111BA  SHARADA VOWEL SIGN VOCALIC L

and

111BB  SHARADA VOWEL SIGN VOCALIC LL

Accepted

TE.2. Page 53, Clause 33, Code Charts and lists of character names – Miao

The glyphs for U+16F2C - U+16F32 have been reported as being in error, and the USNB has received confirmation about the errors and the corrected glyphs from Adrian Cheuk. The correct glyphs are shown on the right column.

Proposed change by US:

Correct the glyphs as follows:

 16F2C	 16F30
 16F2D	 16F31
 16F2E	 16F32
 16F2F	

Accepted

See also comment from China about Miao

TE.3. Page 53, Clause 33, Code Charts and lists of character names – Tamil Supplement

Based on the extensive comments contained in UTC document [L2/16-039](#), the US considers Tamil Supplement not yet mature enough to progress to an enquiry ballot.

Proposed change by US:

Retain Tamil Supplement at committee draft level. If te.3, te.5, and te.6 are accommodated, the USNB will change its vote to yes.

Accepted in principle

See also comment T. from India. However, the Tamil Supplement will be removed from the 5th edition addition but will not be added in any committee draft level ballot until there is a better consensus.

TE.4. Page 53, Clause 33, Code Charts and lists of character names – Soyombo

The glyph for U+11A98 SOYOMBO GEMINATION MARK is incorrect; the triangle should be directly above the dotted circle, not above and to the right.

Proposed change by US:

Correct the glyph by centering the triangle above the dotted circle.

Accepted

TE.5. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2D1D1

An error has been identified in CJK Extension F for the glyph and source of U+2D1D1.

Proposed change by US:

Move the glyph and source KC-06578 from U+2D1D1 (=JMJ-059415) to U+2E068 (=JMJ-059958). If te.3, te.5, and te.6 are accommodated, the USNB will change its vote to yes.

Accepted

See also comments CJK.2 from China, T2 from Japan, T7 and T8 from UK.

TE.6. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2D394

A second error in CJK Extension F has been found: the source and glyph U+2D394 is incorrect.

Proposed change by US:

Move the glyph and source KC-07044 from U+2D394 (=USAT-03073) to U+2DA7C (=GZJW-01933). If te.3, te.5, and te.6 are accommodated, the USNB will change its vote to yes.

Accepted

See also comments CJK.2 from China, T2 from Japan, T7 and T8 from UK.

Editorial comments:

E.1. Page 53, Clause 33, Code Charts and lists of character names – Bengali

The header above U+09FC is spelled “SignS”.

Proposed change by US:

Make the final “S” lowercase

Accepted

E.2. Page 53, Clause 33, Code Charts and lists of character names – Ideographic Symbols and Punctuation

The heading above U+16FE1 is “Nushu mark”, but the “u” should contain an umlaut

Proposed change by US:

Change “Nushu” to “Nüshu”.

Accepted

E.3. Page 53, Clause 33, Code Charts and lists of character names – Nushu

The annotation for U+1B1FE is spelled “Nushu”, but the “u” should contain an umlaut.

Proposed change by US:

Change “Nushu” to “Nüshu”.

Accepted

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: ISO/IEC JTC1/SC2/WG2/IRG
Title: UK Response to Comments on CJK Ext. F (IRGN2146)
Source: Andrew West
Status: Member Body Contribution
Action: For consideration by IRG
Date: 2016-05-21

This document provides the UK responses to comments on CJK Ext. F in ISO/IEC 10646 CD.2 (see IRGN2146).

2D0D6 and 52BD

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

古壮字字典 page 284

𠄎 (勒、烈、捌)
lieb[li:p⁵] 破(竹篾):
~ 笑孛打繡。Lieb ruk
hag dajsan. 破篾学编织。

中华字海 page 195

𠄎 liè 音列] 有力。见《集韵》。

Just because two radicals (刀 and 力) in this case) are similar in shape does not mean that characters with these radicals should be unified. Unification should not be based simply on whether two characters look similar, but should take into account the meaning and usage of the two characters.

2D0E9 and 2B9AF

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

古壮字字典 page 406

斫 (斫、搏、斫、斫、斫、斫、斫)
raemj [vam'] 斫: ~
斫。 raemj faex. 斫树。

汉语方言大词典 page 4647

斫 *lè 斫
【斫瓜】 <名> 黄瓜。客话。江西上犹社溪 [le⁵⁵ kuə²⁴]。

Just because two radicals (刀 and 力) in this case) are similar in shape does not mean that characters with these radicals should be unified. Unification should not be based simply on whether two characters look similar, but should take into account the meaning and usage of the two characters.

2D421 and 225BB

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

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.

2D81D and 22AE4

2E82F and 2882E

The components 凵 and 凵 have different meanings and are not unifiable. Therefore 2D81D should not be unified with 22AE4, and 2E82F should not be unified with 2882E.

2DA99 and 6840
2E1D3 and 25E4F

The components 𠂇 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.

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.

2E0DD and 2581E

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

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

2DBF6 and 6C13
2E295 and 262DF
2E330 and 267C0

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

2E1CF and 7C87
2E9E0 and 29273

In many cases 𠂇 is a variant of 𠂇, but the difference is quite significant, and it is not clear that characters with these two components should be unified. In Ext. F there are 83 characters with the 𠂇 component, and it is not clear why the two characters 2E1CF and 2E9E0 have been singled out for unification.

A few of many examples of characters with 𠂇 component in Ext. F which may be variants of other characters include: 2DECD (JMJ-058092) 𠂇𠂇𠂇𠂇 may be a variant of 7557 𠂇; 2EAFA (JMJ-059086) 𠂇𠂇𠂇𠂇 and 2EAFB (JMJ-059089) 𠂇𠂇𠂇𠂇 may both be variants of 9B5A 𠂇; 2E678 (JMJ-060201) 𠂇𠂇𠂇 may be a variant of 27D28 𠂇; and 2D023 (JMJ-

056971) may be a variant of 5154 兔. Given that these characters are not proposed for unification, we suggest that 2E1CF and 2E9E0 should remain in Ext. F.

2E484 and 2E5AB

There is no precedence for unifying 壯 and 壯, which in China are treated as traditional and simplified forms rather than glyph variants. Therefore 2E484 should not be unified with 2E5AB.

2E5E8 and 2789B

It is not clear that these two characters are cognate, and should be unified. 2E484 consists of three 見 side-by-side, but 2E5AB is unrelated to 見, and is probably a corrupt form of 7516 𠄎. Unless evidence can be produced to show that 2E5E8 and 2789B are cognate they should not be unified.

Feedback to ISO/IEC JTC1/SC2/WG2/IRG N2146
Japan Position to ISO/IEC 10646 5th CD2 Ballot Comment

1. Ballot comments on J-source Characters from other members

#	UCS	WG2 N4713	Japan position
1 a)	2CF16 乙 5.1 此 JMJ-057888 2CF16	Propose accepted 5.1 -> 5.3	Japan agrees with changing Rad.SC (5.1) to 5.3.
1 b)	2CF25 乙 5.6 𠄎 𠄎 JMJ-057893 2CF25	Propose accepted 5.6 -> 5.8	Japan agrees with changing Rad.SC (5.6) to 5.8.
1 f)	2D170 卜 25.5 𠄎 JMJ-058577 2D170	Propose accepted 25.5 -> 25.4	Japan agrees with changing Rad.SC (25.5) to 25.4.
1 p)	2DA5F 木 75.3 𠄎 JMJ-059709 2DA5F	Propose accepted 75.3 -> 75.4	Japan agrees with changing Rad.SC (75.3) to 75.4.
1 ae)	2E37A 白 134.10 興 JMJ-060094 2E37A	Propose accepted in principle 134.10 -> 134.9	Japan agrees with changing Rad.SC (134.10) to 134.9. The radical component of 2E37A is outer component.
1 af)	2E38A 舛 136.5 𠄎 𠄎 JMJ-058426 2E38A	Propose accepted 136.5 -> 136.4	Japan agrees with changing Rad.SC (136.5) to 136.4.

#	UCS	WG2 N4713	Japan position
1 al)	<p>2E822 酉 164.2 𩇛 JMJ-058883</p> <p>2E822</p>	<p>Propose accepted in principle 164.2 -> 164.3</p>	<p>Japan agrees with changing Rad.SC (164.2) to 164.3 2E820 Rad.SC is 164.1, hence Rad.SC of 2E822 should be 164.3.</p> <p>2E820 酉 164.1 𩇛 JMJ-058880</p>
3.5	<p>2E41B 艸 140.10 𩇛 JMJ-022574</p> <p>2E41B</p>	IRG discussion	<p>Japan does not accept proposal. The glyph shape of JMJ-022574 is appropriately designed in accordance with a design policy of IPAmjMincho Font. See the following example:</p> <p>走 徒 MJ025169 MJ011311 U+8D70 U+5F92</p>
3.13	<p>2DF02 疋 103.4 𩇛 JMJ-058113</p> <p>2DF02</p>	IRG discussion	<p>Japan does not accept proposal. The glyph of JMJ-058113 is appropriately and intentionally designed in accordance with the glyph shape of evidence (Shin-Daijiten #10247).</p> <p>10247 𩇛 10246 𩇛 の俗字</p>

#	UCS	WG2 N4713	Japan position
3.15	2D16F 卜 25.5  JMJ-058579 2D16F	IRG discussion	The glyph shape of JMJ-058579 is appropriately designed in accordance with a design policy of IPAmjMincho Font. 2D16F 卜 25.5  JMJ-058579 2D170 卜 25.5  JMJ-058577 14490  14486  の俗字

2. Ballot comments from Japan

Following to WG2 M64.11 “WG2 recommends that IRG reviews its CJK unification rules to minimize the number of glyph variants that are coded as separate characters.”, Japan proposed the unifications for 48 CJK F characters.

Out of those characters, #19 and #20 below are closely related to the comment by UK.

Japan commented the characters on left column and the characters on right column should be unified.

19	2DB0F 木 75.12  KC-07047	69FB 木 75.11      G3-416A HB2-E5E2 T2-512B J0-4450 K0-5027
20	2DB54 欠 76.8  JMJ-057882	238C7 欠 76.7   UCS2003 GHZ-80027.22

On the other hand, UK commented the following two characters should not be unified.

CJK F	Comment

<p>2E1F5 米 119.9</p> <p>𪗇 𪗇</p> <p>JMJ-058296 USAT-00062</p>	<p>U+2E1F5. The two source glyphs (JMJ-058296 and USAT-00062) have different left hand components, which would not seem to be unifiable. Compare U+2DB74 (USAT-00811) and U+2E00D (USAT-02508) which have the same difference in left hand component, but are not unified.</p>
--	--

The decision on whether these two characters are unified or not unified should be consistent. Therefore, if IRG experts reach the consensus not to unify JMJ-058296 and USAT-00062, Japan withdraws the #19 and #20.

(end of document)

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N2146

Date: 2016-05-22

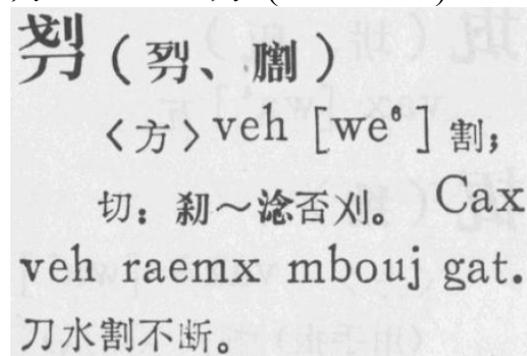
Source:	China
Meeting:	IRG#46, Beijing, China
Title:	Responses to Draft disposition of comments on CJK_F in ISO/IEC CD.2 10646 5th edition
Status:	Member's submission
Actions required:	FYI
Distribution:	IRG
Medium:	Electronic
Pages:	2

We basically accept the proposed dispositions by the UCS Project Editor in IRGN2146, and will accept the conclusion to be made at IRG#46. But the 2 proposed unifications on page 18 (Japan comment T1) of IRGN2146 are not accepted by now:

(1) U+2D0D6 𠄎 vs U+052BD 𠄎

2	2D0D6 刀 18.6 𠄎 GZ-4861102	52BD 力 19.6 𠄎 GE-2353	𠄎	𠄎	𠄎
			T4-2550	J1-3364	K2-2532

The evidence of CJK_F one (U+2D0D2) shows that it belongs to radical 刀 rather than 力 (U+052BD).



(G_Z4861102, CJK_F evidence: sawndip-evdince.pdf page 182)

We can not provide evidence of U+052BD, the source index GE means this G glyph is horizontally derived from T, J, K.

(2) U+2D0E9 𣎵林刀 vs U+2B9AF 𣎵林力

3	2D0E9 刀 18.8 𣎵 GZ-4062201	2B9AF 力 19.8 𣎵 GZFY-00521
---	------------------------------------	------------------------------------

The evidence shows that the CJK_F one (U+2D09E) belongs to radical 刀 rather than力 (U+2B9AF), while the component力 of U+2B9AF is also used as phonetic part.

𣎵 (藤、擗、唸、劔、
斫、斲)
raemj [ɣam³] 斫: ~
斲。 raemj faex. 斫树。

(G_Z4062201, CJK_F evidence: sawndip-evdince.pdf page 186)

𣎵 •力 为力

【斲瓜】<名>黄瓜。客话。江西上犹社溪 [lo²⁴ kua²⁴]。

(GZFY-00521, Hanyu Fangyan Dacidian 汉语方言大辞典)

End

SAT response to the IRGN2146

The following indicates the page numbers and the items of the IRGN2146 document and the SAT responses.

pp. 14-15

- Fonts to be improved:
 - 2CFDF usat-05762
 - 2D102 usat-00875
 - 2DE2F usat-03390
 - 2DFC5 (USAT font to be improved)
- SAT sent the revised fonts according to the indications.

pp. 18-19.

- T1. Page 2489-2582, Clause 33 – “CJK Unified Ideographs Extension F”
- SAT follows decision by IRG on the proposal.

p. 21

- E5. Page 2489 – 2582 Clause 33 “CJK Unified Ideographs Extension F”
- SAT fixed the overlapping problem and sent new fonts to the editor.

p. 25

- T6. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2D0F0
- SAT basically agrees to unify, but follows decision by IRG.

p. 25

- T9. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2DDC6
- SAT basically agrees to unify, but follows decision by IRG.

p. 25

- T10. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2DF3B
- SAT agrees to this change, if IRG accepts.

p. 26

- T11. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2DFC5
 - SAT revised the font.

p. 26

- T13. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2E1F5
 - SAT basically agrees to unify, but follows decision by IRG.

p. 26

- T14. Page 53, Clause 33, Code Charts and lists of character names – CJK Ext. F – U+2E321
 - SAT basically agrees to unify, but follows decision by IRG.

Additional Feedback to Unifications/Disunifications in IRG N2146

suzuki toshiya, 2016-05-23

In IRG46, a few unifications are regarded to collect more information to make a decision. Here is my individual contribution.

1. Unification Proposed by Japan, #26

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
----	------------------------------------	---

The number of the character including 亾 is 222, and that of the character including 亡 is 371

The number of existing disunifications are 23 pairs:

- | | | | |
|----------------|------------|----------------|------------|
| 1) 亾(U+4EBE) | 亡(U+4EA1) | 13) 纒(U+25FEA) | 纒(U+25FFC) |
| 2) 匄(U+5303) | 匄(U+5304) | 14) 翬(U+263F6) | 翬(U+26400) |
| 3) 諛(U+8A64) | 諛(U+279E2) | 15) 苾(U+26B06) | 苾(U+8292) |
| 4) 罔(U+34BA) | 罔(U+7F54) | 16) 螭(U+27351) | 螭(U+8744) |
| 5) 兪(U+3829) | 兪(U+5DDF) | 17) 覩(U+2782C) | 覩(U+27830) |
| 6) 幌(U+3843) | 幌(U+3846) | 18) 迤(U+2845F) | 迤(U+28451) |
| 7) 惘(U+FA86) | 惘(U+60D8) | 19) 鋹(U+289B5) | 鋹(U+92E9) |
| 8) 囟(U+20548) | 囟(U+20543) | 20) 駮(U+298AF) | 駮(U+298C7) |
| 9) 妄(U+216B6) | 妄(U+5984) | 21) 鵠(U+298C8) | 鵠(U+298D0) |
| 10) 寃(U+219EC) | 寃(U+219FD) | 22) 鯨(U+29DA1) | 鯨(U+29DB6) |
| 11) 忙(U+225C5) | 忙(U+5FD9) | 23) 盞(U+2C417) | 盞(U+25043) |
| 12) 磁(U+254B4) | 磁(U+786D) | | |

Therefore, the unification of this pair would be useful to reduce the codepoints for the variant characters.

2. Unification Proposed by Japan, #27

27	2DEF3 <small>田 102.12</small>  JMJ-058106	2DEF8 <small>田 102.14</small>  JMJ-058108
----	---	---

The numbers of characters including 人 or 口 are more than 10000 (16960 and 18595), and the mechanical replacements could include the non-cognate pairs like “呈” and “全”, so I manually picked following disunification examples. JMJ-057407

2D607
工 48.5


1) JMJ-057407 JMJ-057407 聖(U+2A845)

2D7F2
戈 62.4

 JMJ-057538

2) JMJ-057538 笈(U+229B8)

3) 坐(U+5750) 聖(U+212D1)

4) 墻(U+58BB) 墻(U+2BBB6)

5) 檣(U+6AA3) 檣(U+237B8)

6) 賚(U+8CEB) 賚(U+8CF7)

7) 喪(U+462E) 喪(U+55AA)

8) 晉(U+2B784) 晉(U+23206)

3. 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) (U+2BE18) 葬(U+22348).

In CJK F, following 2 might be similar cases.

2D726 怨
心 61.5
JMJ-057499 =怨

2D729 怨
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)

4. Disunification Proposed by UK (T13)



Following pairs might be considered for the consistency.



Similar disunifications are found in 13 pairs:

- | | | | |
|--------------|-------------|----------------|-------------|
| 1) 𪗇(U+4FE3) | 𪗇(U+3468), | 8) 候(U+20240) | 候(U+5019), |
| 2) 𪗇(U+63C6) | 𪗇(U+22C95), | 9) 𪗇(U+202BE) | 𪗇(U+202F3), |
| 3) 𪗇(U+660A) | 𪗇(U+2AC16), | 10) 𪗇(U+21688) | 𪗇(U+25404), |
| 4) 𪗇(U+6E40) | 𪗇(U+23E8D), | 11) 𪗇(U+294A0) | 𪗇(U+294D9), |
| 5) 𪗇(U+7678) | 𪗇(U+24F29), | 12) 𪗇(U+2A7E9) | 𪗇(U+20915), |
| 6) 𪗇(U+8135) | 𪗇(U+2B786), | 13) (U+2ABF6) | 𪗇(U+65CF). |
| 7) 𪗇(U+8475) | 𪗇(U+26DE1), | | |

However, it should be reminded that IRG had once decided not to deal Ext B as the sufficient rationale to disunify the variants.

(end of document)

IRG N2146, Individual Feedback to UK Response to CD2 ballot comments

suzuki toshiya, 2016-05-23

UK response may induce some unified characters to Ext F; 13 JMJ characters, 1 KA character, 3 USAT characters.

力 and 刀 could not be recognized as the significant shape difference to disunify the cognate characters. See Annex S source separation example:

券 券 TJ
5238 52B5

2D0D6 and 52BD

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

if non-cognate, they should be disunified.

2D0E9 and 2B9AF

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

if non-cognate, they should be disunified.

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 玉 96.13 璦璦璦璦璦 G1-6828 HB1-C0F5 T1-7440 J1-4C5D K2-484F U+74A6	unified by U+74A6, irg40.
--	---	---------------------------

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

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.

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

2D81D and 22AE4 
2E82F and 2882E

The components 𠃉 and 𠃊 have different meanings and are not unifiable. Therefore

if disunified by shape, USAT-01643 and USAT01911 should be backed to Ext F

<p>胸 Z_SAT01643 05972 USAT01643</p>	<p>80F8 肉 130.6 胸 胸 胸 胸 胸 胸 G0-5058 HB1-AFDD T1-5761 J0-363B K0-7D58 V1-637B J+80F8</p>	<p>unified by u+80f8, irg40. </p>
<p>胃 Z_SAT01911 05973 USAT01911</p>	<p>80F7 肉 130.6 胃 胃 胃 GE-3958 T3-3352 K1-7524 J+80F7</p>	<p>unified by u+80f7, irg40.</p>

2DA99 and 6840 
2E1D3 and 25E4F

The components 歹 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

<p>鬱 JMJ-057872 03695 JMJ-057872</p>	<p>2C155 木 75.20 鬱 TE-4826 J+2C155</p>	<p>unified by U+2C155(CJK_E), irg43.</p>
--	--	--

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

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

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

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



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 旻 and 旻 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

侵 JMJ-056916 00250 JMJ-056916	4F07 人 9.4 侵 侵 G5-315B T3-2274 U+4F07	unified by U+4F07, irg40.
般 JMJ-057920 03827 JMJ-057920	6BB7 爻 79.6 般 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 爻(weapon), irg40. irg40.
般 JMJ-057921 03831 JMJ-057921	6BB8 爻 79.7 般 般 般 G5-5065 T4-324E J1-4630 U+6BB8	unified by U+6BB8, postponed changing radical to 爻(weapon), irg40.
般 JMJ-058430 06106 JMJ-058430	822C 舟 137.4 般 般 般 般 般 般 般 G0-3063 HB1-AFEB T1-578F J0-484C K0-5A75 V1-644B U+822C	unified to U+822C, irg41.
設 JMJ-058742 06826 JMJ-058742	8A2D 音 149.4 設 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.

2DBF6 and 6C13

2E295 and 262DF 

2E330 and 267C0

The components 亠 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 荒 G0-3B44 HB1-AFEE T1-5772 J0-3953 K0-7C58 V1-647E U+8352	unified to U+8352, irg41.
--	---	---------------------------

2E1CF and 7C87 
2E9E0 and 29273

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

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

𠂇 JMJ-006465 00492 JMJ-006465	4EA2 → 8.2 𠂇 𠂇 𠂇 𠂇 𠂇 𠂇 G0-3F3A HB1-A4AE T1-446D J0-5036 K0-7971 V1-4A5B U+4EA2	unified to U+4EA2, but not agreed to be a general example, change to R16.0(几), SC=2 and FS=3, irg42. postponed being unified by U+4EA2, postponed changing Radical to 几(table), irg40.
--	--	--

2E484 and 2E5AB 

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

According to SCS, they should be unified

壯 𠂇 GTJ
58EE 58EF

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

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