ISO/IEC JTC1/SC2/WG2 N 4727 Date: 2016-05-27

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

Doc. Type:	Disposition of comments
Title:	Disposition of comments on ISO/IEC CD.2 10646 5 th edition
Source: Project: Status: Date: Distribution Reference:	Michel Suignard (project editor) JTC1.02.10646.00.00.00.05 Distribution WG2 and SC2 2016-05-27 : WG2 SC2 N4446 N4454, various IRG documents
weatum:	Paper, PDF lile

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 <u>Underlined Bold Serif text</u>, *with explanatory text in italicized serif*.

The comments received were 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

The CJK modifications are extensive and described in the following page. These were discussed and agreed by the IRG members during the IRG meeting #46 (Beijing, May 23-27 2016).

Based on these dispositions, the Project Editor and the SC2 secretariat have determined that there is enough consensus to create an enquiry ballot (DIS) for the 5th edition of ISO/IEC 10646.

CJK Changes:

(all code points highlighted in yellow are part of the second part of the CJK extension F that was added later in the process)

RS value change

Current Source		Current	New
Code reference		RS	RS
point		value	value
<mark>2CF16</mark>	JMJ-057888	5.1	5.3
<mark>2CF25</mark>	JMJ-057893	5.6	5.8
2CF36	USAT-00615	5.12	102.9
2D108	USAT-01853	18.11	18.12
<mark>2D170</mark>	JMJ-058577	25.5	25.4
2D2B3	USAT-05759	30.12	30.13
2D2E0	USAT-03237	30.14	30.15
2D332	USAT-00433	30.18	30.19
2D4E9	USAT-04051	40.6	40.5
2D4F2	USAT-02978	40.7	40.6
2D68D	KC-06230	53.17	53.18
2D8FC	USAT-05561	66.9	66.10
<mark>2D9A9</mark>	USAT-03442	72.8	72.9
2DA11	GZ-2501101	72.17	72.16
<mark>2DA1E</mark>	USAT-04132	73.7	72.7
2DA5F	JMJ-059709	75.3	75.4
2DB83	KC-05216	77.24	77.23
<mark>2DB88</mark>	USAT-03735	78.3	78.2
2DCF5	USAT-04005	85.17	85.16
2DDE5	USAT-03103	93.7	93.6
<mark>2DE95</mark>	USAT-05286	96.16	96.17
<mark>2DE9E</mark>	GZYS-00321	97.12	97.13
<mark>2DF35</mark>	KC-05441	104.14	104.12
<mark>2DF3B</mark>	USAT-04032	104.16	61.17
2DFDE	KC-06555	109.11	109.12
2E048	JMJ-058197	112.12	86.13
2E0CD	KC-06592	115.8	115.9
2E0E1	KC-05509	115.10	115.11

Current	Source	Current	New	
Code	reference	RS	RS	
point		value	value	
2E0F9	USAT-80094	115.15	115.16	
<mark>2E113</mark>	GCY-2961.00	116.10	116.11	
<mark>2E172</mark>	USAT-04888	118.9	118.8	
<mark>2E207</mark>	USAT-02077	119.14	119.13	
<mark>2E28B</mark>	KC-05582	122.5	122.4	
2E2B2	GPGLG-3035	123.7	123.6	
<mark>2E37A</mark>	JMJ-060094	134.10	134.9	
<mark>2E38A</mark>	JMJ-058426	136.5	136.4	
<mark>2E3A7</mark>	USAT-03713	140.4	140.5	
<mark>2E495</mark>	KC-06735	140.14	140.15	
<mark>2E4A3</mark>	USAT-06293	140.15	140.16	
<mark>2E4F1</mark>	KC-06741	141.10	141.11	
<mark>2E4F4</mark>	USAT-80110	141.10	141.9	
<mark>2E573</mark>	GZ-4841302	145.4	145.5	
<mark>2E5CD</mark>	USAT-01060	146.8	146.9	
<mark>2E685</mark>	KC-03956	154.5	154.4	
<mark>2E713</mark>	USAT-01708	157.10	157.11	
<mark>2E7AE</mark>	USAT-03996	162.8	162.9	
<mark>2E822</mark>	JMJ-058883	164.2	164.3	
<mark>2E84B</mark>	KC-06837	164.11	164.12	
<mark>2E8E4</mark>	KC-05832	167.14	167.12	
<mark>2E914</mark>	USAT-03837	169.5	169.4	
<mark>2E9A7</mark>	KC-04405	173.9	173.10	
<mark>2EA47</mark>	KC-05898	184.6	184.5	
2EA85	KC-07176	187.5	187.4	
2EBB3	KC-06965	207.5	207.6	
2EBB6	KC-07188	207.11	207.12	

Notable glyph change:

2DFC5 (radical change)

Character deletions (24):

Code	Unified	Deleted source reference
point	with	(Font code)
2D2E5	210A0	USAT-05290 (F1-8E7D)
2D4E5	2D4E9	JMJ-057329 (F1-9063)
2D584	21D45	USAT-00821 (F1-90EB)
2D8B4	6503	USAT-01809 (F1-93E3)
<mark>2D93A</mark>	2B780	JMJ-057640 (F2-3104)
2D99C	666B	JMJ-059691 (F1-94AE)
2DA4C	2339F	JMJ-037911 (F2-3412)
2DA99	6840	JMJ-057740 (F2-3480)
2DAE0	23624	USAT-02858 (F1-95DC)

Code	Unified	Deleted source reference
point	with	(Font code)
2DAFD	6A7B	USAT-01188 (F1-95F7)
2DB74	2E00D	USAT-00811 (F1-965B)
2DB8C	239CA	USAT-05410 (F1-9672)
<mark>2E100</mark>	4193	USAT-02971 (F2-5366)
<mark>2E168</mark>	25C4D	USAT-60250 (F2-5473)
2E1CF	7C87	USAT-03923 (F2-5604)
2E1D3	25E4F	JMJ-058287 (F1-98ED)
2E341	8188	JMJ-058398 (F1-9908)
2E368	2693C	JMJ-060086 (F1-990B)

Code point	Unified with	Deleted source reference (Font code)
2E40B	8521	JMJ-060130 (F1-991C)
<mark>2E484</mark>	2E5AB	USAT-03743 (F2-6355)
<mark>2E5E8</mark>	2789B	JMJ-058726 (F2-6788)

Code point	Unified with	Deleted source reference (Font code)		
<mark>2E615</mark>	8AAA	JMJ-058743 (F2-6838)		
<mark>2E63A</mark>	27AD5	USAT-00800 (F2-6879)		
<mark>2E9E0</mark>	29273	USAT-03573 (F2-7910)		

Concerning 2D4E5-2D4E9, JMJ-057329 (F1-9063) is moved to 2D4E9, JMJ-057331 (F1-9067) is removed from 2D4E9, the RS value is updated from 40.6 to 40.5.

Concerning 2E484-2E5AB, USAT-03743 (F2-6355) is moved to 2E5AB with KC-06763 (F2-6733), font remap necessary.

Glyph/source reference moves (2):

KC-06578 to U+2E068 (with JMJ-059958) KC-07044 to U+2DA7C (with GZJW-01933)

Character additions (12):

JMJ-057583, JMJ-058052, JMJ-059628, USAT-01643, USAT-01911, JMJ-057168, JMJ-056916, JMJ-057920, JMJ-057921, JMJ-058430, JMJ-058742, JMJ-058447

Please refer to following pages for details.

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, w2ith 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 (かかか), but this scheme mentions none; the final forms of traditional Mongolian letter NA have at least 6 forms (盛盪丙 磕 礢 礙), but this scheme has 2 forms (盛盪) 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 *, *; * and * are individual letters and no need to be differentiated by FVSs. This scheme specifies that * and *" do not need FVSs while * and * 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 монголжин 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 * eletters!

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.

<u>Noted</u>

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.

<u>Noted</u>

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 $\rightarrow \mu/\mu$) and incorrect (1824 $\rightarrow V[u]$).

<u>Noted</u>

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, OSs 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. Page 6

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.

<u>Noted</u>

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 'X' in the name list for both characters as in:
 - 19AA స్రNEW TAI LUE LETTER HIGH SUA※ NEW TAI LUE LETTER HIGH SVA19AB ရွNEW 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 m 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 6 NEW TAI LUE LETTER HIGH SUA

Technical comment on Miao

The following glyphs should be corrected as below:

- 16F2C **E** MIAO LETTER NYA
- 16F2D € MIAO LETTER NYHA
- 16F2E C MIAO LETTER TSHA
- 16F30 C MIAO LETTER YI TSHA
- 16F32 I MIAO LETTER REFORMED TSHA
- 16F31 C MIAO LETTER YI DZHA (Glyph position correct; no need to change)

Explanations:

- 16F2C $\boldsymbol{\epsilon}$: This character is used in the *Normalised* but not the *Traditional Miao Orthography*. Moreover, the glyph shape per Miao scholars' design is $\boldsymbol{\epsilon}$.
- 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 ϵ .

- 16F2E ^C: 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 ^C: 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 \bot : 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: $16F2C \rightarrow 16F32$ $16F2D \rightarrow 16F2C$ $16F2E \rightarrow 16F2D$ $16F2F \rightarrow 16F2E$ $16F30 \rightarrow 16F2F$ 16F31 unchanged $16F32 \rightarrow 16F30$ 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): <u>http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg45/IRGN2130_HKSAR_Review.pdf</u>, 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 <u>http://appsrv.cse.cuhk.edu.hk/~irg/index.htm</u>.)

1. Radicals and Stroke Counts

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



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



5.6 -> 5.8

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

2CF36 Z 5.12 USAT-00615

Accepted in principle

5.12 -> 102.9. The radical (102: \boxplus) change and resulting value is preferred over 5.12 -> 5.13.

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

2D02F 儿 10.11 GZ-0062201

Not accepted

It would be 10.11 -> 10.10. Rationale to keep current residual stroke count is provided in IRG N954AR #78

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

2D108 刀 18.11 USAT-01853 Accepted

18.11 -> 18.12

f) 2D170, SC 4



Accepted 25.5 -> 25.4

g) 2D2E0, SC 15, reference 6A06, 7483, and 96E2 2D2E0 _{□ 30.14}

USAT-03237

Accepted

30.14 -> 30.15

h) 2D332, SC 19 2D332 口 30.18 USAT-00433

Accepted

30.18 -> 30.19

i) 2D4F2, SC 6, the SC of the residual component is 6 in KangXi dictionary, reference 2DD23, 2DA99 $^{2D4F2}_{\sim 40.7}$

7 **7** USAT-02978

Accepted

 $\overline{40.7} \rightarrow 40.6$

j) 2D68D, SC 18, reference 50B2, 71AC, and 8D05 2D68D 广 53.17 <u>KC-06230</u> <u>Accepted</u> 53.17 -> 53.18

k) 2D8A8, SC 12, reference 2E282 and 2E8B6 2D8A8 $\neq 64.13$

USAT-05373 Not accepted

It would be 64.13 -> 64.12. Rationale to keep current residual stroke count is provided in IRG N954AR #36.

1) 2D8FC, SC 10 2D8FC 支 66.9 USAT-05561

Accepted 66.9 -> 66.10

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

2D9A9 ☐ 72.8 USAT-03442

Accepted

72.8 -> 72.9

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

USAT-04132 Accepted

 $73.7 \rightarrow 72.7$ (radical changed, no change in SC)

o) 2DA45, SC 9 2DA45 _{月 74.10} USAT-04120

Not accepted

It would be 74.10 -> 74.9. Rationale to keep current residual stroke count is provided in IRG N954AR #35.

p) 2DA5F, SC 4, reference 4F3C, 59D2, and 62DF $2DA5F_{\pm 75.3}$ 3

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 _几余

2DB83 止 77.24 近先 KC-05216 <u>Accepted</u>

77.24 -> 77.23

r) 2DB88, SC 2, SC of left component (radical) is 5 in KangXi dictionary, reference 2DB85, 400E, and 5DE7 2DB88 \overline{p} 78.3

F 78.3 ソリ USAT-03735 <u>Accepted</u>

 $\frac{Accepted}{78.3} \rightarrow 78.2$

s) 2DCF5, SC 16, SC of lower right component is 6 in KangXi dictionary, reference 2DD23 2DCF5 ^{水 85.17} (好

USAT-04005

Accepted

85.17 -> 85.16

2DDE5 将 # 93.7 化 USAT-03103

Accepted

93.7 -> 93.6

u) 2D49E, SC 13 [probably a typo for 2DE9E] 2D49E _{女 38.12} USAT-01951

Not accepted

This would be 38.12 -> 38.13. That change was not in HKSAR feedback and is out of sequence, it is probably a typo for 2DE9E which is part of the HKSAR feedback and is not part of the Chinese comments. See T1.az.

v) 2DF35, SC 12 2DF35 104.14 KC-05441

<u>Accepted</u> 104.14 -> 104.12. Another opinion was 104.13.

w) 2DFDE, SC 12 2DFDE 目 109.11 日年 KC-06555

<u>Accepted</u> 109.11 -> 109.12

x) 2E0E1, SC 11, reference 5368, 21A9D, and 25801 2E0E1 禾 115.10 KC-05509

Accepted 115.10 -> 115.11

y) 2E0F9, SC 16, reference 417B, 6A06, and 7483 2E0F9 天京 [★] 115.15 USAT-80094

Accepted

115.15 -> 115.16

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



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 なな

2E172 竹 118.9 USAT-04888

<u>Accepted</u> 118.9 -> 118.8

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



119.14 -> 119.13

ac) 2E28B, SC 4, reference 7F52

<u>Accepted</u> 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 ¥.

¥ 123.7 GPGLG-3035 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 134.10

E 134.10 <u>JMJ-060094</u> <u>Accepted in principle</u> 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.



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



 $140.4 \rightarrow 140.5$



141.10 -> 141.11

ai) 2E4F4, SC 9, same rationale as for 2E38A 2E4F4 # 141.10

USAT-80110 Accepted

141.10 -> 141.9

aj) 2E685, SC 4, same rationale as for 2E38A 2E685154.5

<u>KC-03956</u> <u>Accepted</u> 154.5 -> 154.4

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



<u>Accepted</u> 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



Accepted in principle 164.2 -> 164.3

am) 2E84B, SC 12 2E84B 164.11 KC-06837 Accepted 164.11 -> 164.12

an) 2E8E4, SC 12 2E8E4 \$\prod_{167.14}\$

Accepted

167.14 -> 167.12. Another possibility was 167.13.

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

門 169.5 USAT-03837

<u>Accepted</u> 169.5 -> 169.4

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



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 食 184.6 KC-05898

<u>Accepted</u> 184.6 -> 184.5

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

2EA85 馬 187.5 既 KC-07176

<u>Accepted</u> 187.5 -> 187.4

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



No change: - (2) + - (1) + \boxplus (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 ^{鼓 207.5} 北 広 を た KC-06965

Accepted in principle

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 ^{鼓 207.11} 立 KC-07188

Accepted in principle 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?



Not accepted

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

aw) 2D2B3, SC =13 2D2B3 □ 30.12 USAT-05759 Accepted 30.12 -> 30.13 ax) 2DA11, SC=16?, reference 2A813, 2B249, 4B1D, and 4D43 2DA11 日共 □ 72.17 USAT-05759 GZ-2501101

Accepted

72.17 -> 72.16

ay) 2DE95, SC =17 2DE95 素 96.16 USAT-05286

Accepted

96.16 -> 96.17 (or the two vertical strokes should be merged into one).

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

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



<u>Accepted</u> 115.8 -> 115.9

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



bc) 2E4A3, SC=16 2E4A3 ^{艸 140.15} USAT-06293

<u>Accepted</u> 140.15 -> 140.16

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

bf) 2E7AE, SC=9 2E7AE [№] 162.8 USAT-03996 Accepted

162.8 -> 162.9

bg) 2E9A7, SC=10 2E9A7 ^{〒173.9} 子文

KC-04405 Accepted

173.9 -> 173.10

2. Wrong Positions

KC-06578, $\mathring{\uparrow}$ \diamondsuit 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 **Partially accepted**

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).

2D16F, 2DF02, and 2E41B are intentional and will not be modified. 2D0CF, 2D0D0, 2D171, 2DA69, 2DA90 have one stroke split into 2 strokes. Others are glyph path improvements.

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

2DFC5 目 109.8	腏	畷
	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 $m_{\rm T}$ tamil sign muuvulakku" has to be removed.

Accepted

IN.3. Page 145, Tamil 0BA4

Annotation of OBA4 has reference to 11FD7 which is proposed to be deferred from encoding. See "Proposed change" of Comment "IN 1". <u>Proposed change by India.</u> The annotation "→ 11FD7 ரூ) tamil sign mukkuruni" has to be removed.

Accepted

IN.4. Page 145, Tamil 0BB3

Annotation of OBB3 has reference to 11FD7 which is proposed to be deferred from encoding. See "Proposed change" of Comment "IN 1". <u>Proposed change by India.</u> The annotation "→ 11FD7 IFD 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 Supplemet block which is proposed to be deferred from encoding. See "Proposed change" of Comment "IN 1". <u>Proposed change by India</u> 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 Supplemet block which is proposed to be deferred from encoding. See "Proposed change" of Comment "IN 1". <u>Proposed change by India</u>

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

IN.7. Page 145, Tamil 0BFA

Annotation of OBB3 has reference to 11FF1 which is proposed to be deferred from encoding. See "Proposed change" of Comment "IN 1". <u>Proposed change by India.</u>

The annotation " \rightarrow 11FF1 \mathbb{m} 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	2DAF0 + 7/2 23624 + 4/7 + 4/7 + 4/7	
1 2D0D2 USAT 0005	B9	17 * 75.10 个示 * 75.11 个宗 个示 * 75.11 US2703 GKX-0549.25 T44A3F	
2 2 刀 18.6 子 J 18.6 子 J 18.6	^D 列 列 列 列	18 本 75.11 USAT-01188 6678 示 75.12 GK-6859 k2	-3D3A
GZ-4861102 2D0E9 赤木 2B9A 3 刀 18.8 プリ わり	GE-2353 T4-2550 J1-3384 K2-2532 AF 六 9.8 六	19 ^{2DB0F} 槻 ^{69FB} 桃 根 根 根 根 482-552 12578 49450 465	月 027
GZ-4062201 2D10E 教 2D13 4 刀 18.11 外 2D13 力 19:	<u>GZFY-00521</u> 3D	20 2DB54	
USAT-05989 2D136	USAT-01966 78 齐文 齐文 8.9 子文	21 2DB8C <u></u>	
USAT-60035 6 2D2E5 IIズ 210A □ 30.14 IIズ 30.1	ucs2003 GHZ-8001024 A0	22 20 20 20 20 20 20 20 20 20 20 20 20 2	
USAT-05290 7 2D421 安 225日 7 久 35.5 安 心 61.3	ucs2003 gKX-0213.07 T45A79 BB 发发发发	23 2797 23 2797 2797 279.6 日又 日又 55.603 13.216B	
USAT-02739 8 2D44B 大 37.6 大臣 77E9 失 111	9 115 矩矩矩矩矩矩矩	24 2DBC0 展史 ^{(2, 79:10} 展史 _{(2, 79:10} 局段 _{(2, 79:10} (2, 79:10) (2,	股 殿
9 2D584 11 46.3	GU-3EDS HBT-AF/8 11-36/8 JU-3668 KU-4F-36 V1-6068 45	25 2DBC2 279.11 <u>期</u> <u>JMA 057928</u> 6BC4 <u></u> 文 79.10 <u></u> <u>BC</u> <u></u> <u></u> <u>BC</u> <u></u> <u>BC</u> <u></u> <u></u> <u>BC</u> <u></u> <u></u> <u>BC</u> <u></u> <u></u> <u>BC</u> <u></u> <u></u> <u>BC</u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u>	几 ズ 830
10 ² 206E1	94 役 役 役 役 役 役 00555 HELATO TLANS	26 20BF6 低低 GC13 正 正 正 正 正 U U U U U U U U U U U U U U	民 氓 86C ¥1-5858
11 2D818	5.4 投設 L <thl< th=""> <thl< th=""> <thl< th=""></thl<></thl<></thl<>	27 2DEF3 III 102.12 JAL-058106 2DEF8 III 102.14 JAL-058106 JAL-058108	
12 2D81D 手 64.5 USAT-05471 22AE 手 64.9	E4 拘 拘 拘	28 年 115.9 年	
13 2D8B4 读 6503 13 ^{手 64.14} 读 564.14 USAT-01809	³ 擦 擦 擦 擦 擦 擦 。 擦 。 \$23746	29 ^{2E100}	
14 2D99C 正存 日 72.8 正存 JMJ-059691 6666B	B	30 年168 新 11 118:9 新 USAT-80250 25C4D 新 前 118:10 新 UCS2003 T5-5658	
15 2DA4C 月 74.12 JMJ-037911 2339 月 74.	9F 時 時 ^{UCS2003} 15-546D	31 2E1CF 航 7C87 米 1194 航 航 航 称 称	Г 146
16 2DA99 获 6840 * 75.7 承 75.1	0 56 架架架架架架架 60.666E HB1.4EE5 T1.5629 J0.585C K0.487A V1.5971	32 2E1D3 32 * 119.5 JMJ-058287 25E4F * 119.5 UCS2003 GKX-0508.12 T6-4C63	

33	2E232 _{系 120.9}	紙 USAT-01467		27322 _{虫 142.8}	給	盆 GKX-1088.2	2 15-4			
34	2E295 [9] 122.9	託 JMJ+058343		262DF 阿 122.9	以 UCS2003	区 KX-0948.34	下7-2	575		
35	2E304 耳 128.12	聲 JMJ-058370		8072 耳 128.11	聲 G1-4979	聲 HB1-C16E	設 耳 ^{[1.7478}	<u>臀</u> J0-6661	聲 K0-6122	聲 ^{V1-6363}
36	2E314 肉 130.4		月又 USAT-60289	809E 肉 130.3	<u></u> (E-3952		EX T4-2470	<u>敗</u> J1-556E		
37	2E330 肉 130.10	扇 JMJ-058394		267C0 肉 130.10	<u>肩</u> UCS2003	扉 GHZ-32104.0	4 T6-6	尾 170		
38	2E341 肉 130.12	腪 」MJ-058398		8188 肉 130.10	膈 60-6875	居)	日	膈 J0-672D	膈 K0-4C2E	
39	2E368 至 133.4	室 JMJ-060086		2693C 至 133.5	<u>室</u> UCS2003	<u>声</u> GHZ-42816.0	1 T4-3	11) And 456		
40	2E40B 99 140.10	<u>茶</u> JMJ-060130		8521 ⊯ 140.11	蔡 60-324C	蔡 ^{HB1-BDB2}	<u>获</u>	蔡 J0-6871	蔡 K0-7379	
41	2E484 (#) 140.13	衷 USAT-03743		2E5AB 衣 145.10	衷 KC-06763					
42	2E5E8 _{見 147.14}	月月月 JMJ-058726		2789B ^{見 147.14}	間目 加九 UCS2003	目目目 ルルル GHZ-63676.0	7			
43	2E615 _{言 149.6}	<u>説</u> JMJ-058743		8AAA 育 149.7	說 GE-3D73	說 HB1-BBA1	記 ^{[1-6B29}		說 K0-6063	
44	2E63A 言 149.10	記 USAT-00800		27AD5 育 149.11	談 UCS2003	主奴 日不 GKX-1177.12		マ 下 258		
45	2E752 車 159.5	 更 JMJ-058827		8ED7 車 159.4	軗 GE-3F62	 世 HB2-D7DF	 段 [2-3962			
46	2E769 車 159.9	 助		2833A 車 159.8	翰	翰 H-BECA				
47	2E82F 酉 164.5	酌 KC-05802		2882E 西 164.6	西 UCS2003	函 GHZ-63581.1	т т т 5-5	XI MD		
48	2E9E0 革 177.4	貥		29273 革 177.4	翫	斯 GHZ-74327.2	0			

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+2D884, U+2D99C, U+2DA4C, U+2DA99, U+2DAE0, U+2DAFD, U+2DB0F, U+2DB54, U+2DB8C, U+2DB90, U+2DB88, U+2DBC0, U+2DBC2, U+2DBF6, U+2DEF3, U+2E0DD, U+2E100, U+2E168, U+2E1CF, U+2E1D3, U+2E322, U+2E395, 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.

Partially accepted

Entries 19 and 20 are withdrawn by Japan. Among the rest, unification is accepted for the following entries: 6, 9, 13-18, 21, 29-32, 38-44, and 48.

In addition, for consistency with the unification/dis-unification decisions made above, the following 12 characters are added to CJK ext. F:

Caused by the maintained dis-unification of 2D421-225BB:

瑷 04639	96 王 13	.0 3	- 夏 JMJ-058052
愛	35. ×	0	愛
01644	11	3	JMJ-059628

Caused by the maintained dis-unification of 2D81D-22AE4 (and others):

匈肖

05973







Caused by the maintained dis-unification of 2DEF3-2DEF8 (and others):

	r
Ħ	入
T	マ
010)28



Caused by the maintained dis-unification of 2D0D2-206B9 (and others):



Caused by the maintained dis-unification of 2DBF6-6C13 (and others):



In addition, the following addition was considered in relationship of the maintained dis-unification of 2E0DD-2581E:



as a possible dis-unification from U+26BCC: 26BCC 艸 140.7 GKX-1035.05 UCS2003 T6-4438

It was however decided to keep the unification in that case.

T2. Clause 33 – "CJK Unified Ideographs Extension F"

Following glyphs are placed on same code point by mistake.



Proposed change by Japan.

Ц

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

2E068 示 113.5 JMJ-059958

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



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

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 5^{th} 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] <u>Proposed change by Japan</u> "Note 1", "Note 2" and "Note 3" should be "NOTE 1", "NOTE 2" and "NOTE 3. <u>Accepted</u>

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 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
should be

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.



USAT-00025 <u>Proposed change by Japan.</u> USAT font should be fixed <u>Accepted in principle</u>

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 SUPPLEMEMENTAL 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. <u>Proposed change by Norway.</u>

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 N (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. <u>Proposed change by Norway.</u>

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". <u>Proposed change by U.K.</u> Change "Nushu" to "Nüshu". <u>Accepted</u>

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". <u>Proposed change by U.K.</u> Change to "Source reference file for Nüshu characters". <u>Accepted</u>

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 Page 27 NUSHU CHARACTER-1B2DD * tshew5 should be after NUSHU CHARACTER-1B2DE * lew33 NUSHU CHARACTER-1B2F2 * ku44 should be before NUSHU CHARACTER-1B2F0 * cyu35 <u>Proposed change by U.K.</u> 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.

Not accepted

This was discussed by IRG#46 and rejected. For reference the chart entry looks like:

2D0F0	左 [亍.
刀 18.8	火山	夗

JMJ-059378 USAT-04376

Related to this decision, a new unification was decided among the following proposed code points:

2D4E5	宛	2D4E9	家	宓
40.0	JMJ-057329	40.0	JMJ-057331	USAT-04051

The source JMJ-057331 is removed, and the RS value is 40.5

The following two pairs were also considered:



These are not unified, because in that case the top horizontal stroke makes the difference more significant.

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.

Accepted

For reference the chart entry looks like:



JMJ-057583 was originally proposed with RS: 66.3 (F index 3013):

J.Fr	66.0 支			
12				12
03013	3	4	1	JMJ-0575

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.

Accepted

This was endorsed by IRG#46. For reference the chart entry looks like:

2DF3B	瘛
	USAT-04032

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

Accepted

See also comment CJK T3 from China.

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.

Accepted

This was endorsed by IRG#46. For reference the chart entry looks like:



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 [*ed. Read JMJ-058296*] and USAT-04376 [*ed. Read USAT-00062*] as separate characters, JMJ-058296 under 9 strokes, and USAT-00062 under 11 strokes.

Not accepted

After discussion at IRG#46, the unification was maintained. For reference, the chart entry looks like:



For consistency, the pair U+2DB74-U+2E00D is unified, with resulting RS: 111.11 and the source USAT-00811 removed.

2DB74 2E00D 止乙 止 77.11 天木 矢 111.11 USAT-00811 USAT-02508

In addition, U+2D93A is unified with 2B780 (extension D).



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 [*ed. should be right*] 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.

Not accepted

This was discussed by IRG#46 and rejected. For reference the chart entry looks like:

2E321 肉 130.8



JMJ-058387 USAT-60296

E15. Page 2665, Annex G

Missing space between "and" and "Tangut". <u>Proposed change by U.K.</u> Insert a space between "and" and "Tangut". <u>Accepted</u>

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 <u>L2/15-255</u>, 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:

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:



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 $\underline{L2/16-039}$, 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. <u>Proposed change by US:</u> Correct the glyph by centering the triangle above the dotted circle. <u>Accepted</u>

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. <u>Proposed change by US:</u>

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". <u>Proposed change by US:</u> 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 <u>Proposed change by US:</u> Change "Nushu" to "Nüshu". <u>Accepted</u>

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. <u>Proposed change by US:</u> Change "Nushu" to "Nüshu". Accepted