

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

Doc. Type: Disposition of comments

Title: Disposition of comments on CDAM1 to ISO/IEC 10646 6th edition

Source: Michel Suignard (project editor)

Project: JTC1 02.10646.00.02.00.06

Status: For SC2 consideration

Date: 2022-01-31

Distribution: WG2

Reference: SC2 N4769

Medium: Paper, PDF file

Comments were received from the following members: China, Japan, Korea, 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 has 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 following items reflect the changes to Amendment 1 between CDAM1 and the planned CDAM1.2 based on the ballot disposition:

Removal

UCS Name

11F3C KAWI VOWEL SIGN VOCALIC L
31A86 CJK UNIFIED IDEOGRAPH-31A86

Addition

UCS Name

3222D CJK UNIFIED IDEOGRAPH-3222D

Name changes

UCS Old name

11F42 KAWI SUBJOINER
1F7D9 NINE POINTED STAR

New name

KAWI CONJOINER
NINE POINTED WHITE STAR

Character move

Old UCS New UCS Name

0895 10EFD ARABIC SMALL LOW WORD SAKTA
0896 10EFE ARABIC SMALL LOW WORD QASR
0897 10EFF ARABIC SMALL LOW WORD MADDA

(in addition, all CJK Unified ideographs Extension H between U+31A85 and U+3222D are moved, same total size)

Glyph changes through this ballot (pending receiving font with updated glyphs)

UCS IRG source

317F3 kIRG_VSource VN-F1E36
31F5F kIRG_VSource VN-F1B1F
31F63 kIRG_VSource VN-F085B
31F84 kIRG_VSource VN-F1B10
31FB5 kIRG_VSource VN-F1B17
31FC5 kIRG_VSource VN-F1B19

Annotations, cross reference, source reference added/modified

UCS Name or IRG source

0675 ARABIC LETTER HIGH HAMZA ALEF (preamble modified)
5AB4 kIRG_KSource K6-0004 (horizontal extension)
741F kIRG_KSource K6-0005 (horizontal extension)
74A4 kIRG_KSource K6-0006 (horizontal extension)
74B8 kIRG_KSource K6-0007 (horizontal extension)
7807 kIRG_KSource K6-0008 (horizontal extension)
7A66 kIRG_KSource K6-0009 (horizontal extension)

China: Positive with comments

China votes YES to UCS2021 CDAM1 with comments.

Technical comment:

T1. CJK Unified Ideographs Extension H

One character is to be removed from CJK_H code chart:

U+31A86 (T13-3647) 灑. This character could be unified with U+31A84 (T13-3643) 灑, thus should be removed.

One missing character is to be added back to CJK_H code chart:

閔

The above comments are recorded in IRGN2506 IRG#57 Editorial Report.

Proposed change by China:

Remove T13-6347 and add GZ-4322205.

Accepted

The missing character was the result of a production error in creating the amendment.

See also comments TE.8 and TE.9 from USA.

Editorial comments:

E1. Arabic

The note “Use of these characters is discouraged.” for U+ U+0675 through U+0678 should be removed.

These four letters are not equal to the corresponding target sequences, and the values and orderings are different. This paragraph will mislead CLDR and the input method, and the strings which the end users input and output will be different from the previous. The names of Kazakh compatriots in China always refer to the common Kazakh words written with the Arabic script, these four letters are commonly used. The note “Use of these characters is discouraged.” introduces hidden dangers to the Chinese administration concerning residents and residency.

Proposed change by China:

The note “Use of these characters is discouraged.” for U+ U+0675 through U+0678 should be removed.

Accepted in principle

The note in the code chart says (full text follows):

Use of these characters is discouraged. They were encoded for Kazakh digraphs, but their decompositions do not reflect the preferred order of representation.

There is no dispute about the fact that these digraph letters are in use for Kazakh. The note refers to the encoding of these letters using the precomposed characters. Their decomposition is however anomalous in using the incorrect order, that cannot be changed. Thus, the representation of these letters should use character sequences starting with U+0674, not ending with that code point.

The Unicode Core Specification Chapter 9, page 394 has the following text:

High Hamza. The characters U+0675 arabic letter high hamza alef, U+0676 arabic letter high hamza waw, U+0677 arabic letter u with hamza above, and U+0678 arabic letter high hamza yeh are not recommended for use. Their compatibility decompositions are anomalous: the decomposed sequences are pairs of letters with right-to-left bidirectional character properties, with U+0674 arabic letter high hamza as the second letter. When processed using the Unicode Bidirectional Algorithm, the hamza will appear to the left of the other

letter. Thus, the ordering of characters in the decomposition sequences are the reverse of what is expected. Accordingly, appropriately-ordered pairs of letters beginning with U+0674 arabic letter high hamza should be used. For example, the sequence should be used rather than U+0675 arabic letter high hamza alef.

While the note addresses the first part of the Core Specification it does not provide a solution. To address the concern from China the note in the code chart could be changed as follows:

Use of these characters is discouraged. They were encoded for Kazakh digraphs, but their decompositions do not reflect the preferred order of representation. Accordingly, the representation of these Kazakh digraphs should use the correct order of decomposed sequences.

And the code points 0675..0678 should have the additional annotations:

0675 ARABIC LETTER HIGH HAMZA ALEF

- should be represented using the sequence 0674 0627

0676 ARABIC LETTER HIGH HAMZA WAW

- should be represented using the sequence 0674 0648

0677 ARABIC LETTER U WITH HAMZA ABOVE

- should be represented using the sequence 0674 06C7

0678 ARABIC LETTER HIGH HAMZA YEH

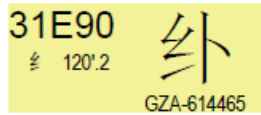
- should be represented using the sequence 0674 0649

Japan: Negative

Technical comment:

TE.1. Clause 34. Code charts... – CJK Unified Ideographs Ext H – 31E90

Some CJK characters in CJK Unified Ideographs Extension H block have the radical number with prime (or dash) above, such as



However, there is no description about this. Section 24.3.1 of ISO/IEC 10646:2020 says “The graphic representation for the radical is shown immediately below the code point, along with the radical number and the stroke count.”.

Proposed change by Japan:

Please specify the description about the radical number with prime.

Accepted

Note that the ‘prime’ notation is described in Table 5 in sub-clause 24.2, but the text in sub-clause in 24.3.1 can be modified to read in a manner similar to sub-clause 24.2:

The graphic representation for the radical is shown immediately below the code point, along with the radical number, optionally followed by an apostrophe for simplified radicals, and the stroke count.

TE.2. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31FC5

Following character is categorized in radical “GLASS”. Please confirm if it is correct.



Proposed change by Japan:

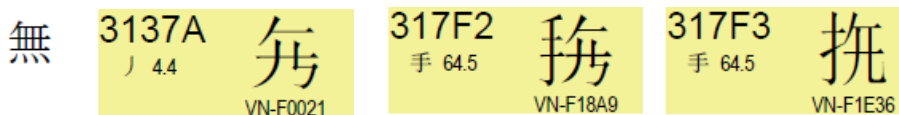
Please move it based on correct radical.

Accepted in principle

The radical is effectively grass, just a different form encoded itself at U+4491 𦰇, so it does not need to be moved. The same radical shape is used for U+31F63 (VN-F085B), U+31F5F (VN-F1B1F), U+31FB5 (VN-F1B17), and U+31F84 (VN-F1B10). The form is more common in manuscripts than print. It could be advisable to change the radical of these 5 characters to use the more traditional form 艸.

TE.3. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 317F3

We think all of 1) U+3137A, 2) right Kanji radical of U317F2 and 3) right Kanji radical of U+317F3 below are normalized form of character “nothing.”



But, only right Kanji radical of U+317F3 is different from others.

Proposed change by Japan:

Please consider changing right Kanji radical of U+317F3 to the one of U+317F2.

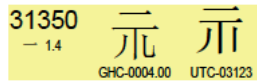
Accepted in principle

What is called right Kanji radical is really the right 'component' because the radical terminology is reserved to the component used for character classification, typically on the left or on the top of the ideograph.

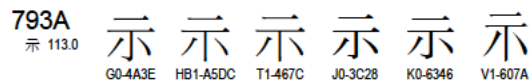
After discussion with experts working on the Vietnam sources, it was agreed to switch the shape of that component for VN-F1E36 to the same as VN-F18A9.

TE.4.1. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31350

Following characters in CJK Unified Ideograph Extension H should be unified with the existing characters in ISO/IEC 10646.



U+31350 above should be unified with U+793A in next page.



Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

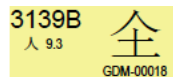
Not Accepted

Most of the concerns shown in the Japanese comments have been addressed by IRG (Ideographic Research Group) in various documents such as the attribute file of WS2017 7.0 (attachment to IRGN2513)

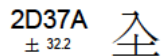
<https://appsrv.cse.cuhk.edu.hk/~irg/irg/irg57/IRGN2513WS2017Attributes.xlsx> and previous versions of that attribute file which have been available to IRG experts. In addition, IRG members use the tool located at <https://hc.jsecs.org/irg/ws2017/app/> to document decisions related to the WS2017 set base for CJK Ext H. The information provided by the tool contains conclusion notes that are duplicated in these dispositions as appropriate. In addition, there are many discussions points captured in these pages explaining the current conclusions. It is likely that an examination of these elements would have addressed concerns from the Japanese NB. Typically, IRG is the forum where these discussions are held, and the result is then transmitted to the ISO/IEC 10646 project editor for further processing. It is unfortunate that Japan experts limited their participation in recent IRG sessions.

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=00002> latest note says: not unified to U+793A 示., IRG 53

TE.4.2. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 3139B



U+3139B above should be unified with U+2D37A below.



Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=00094> latest note says: not unified to U+2D37A for non cognate, 2020-11.

TE.4.3. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31470



U+31470 above should be unified with U+22016 below.

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider

22016
工 48.2 𠄎 𠄎 𠄎
UCS2003 GBK UTC-00069

registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=00356> latest note says:
moved to ws2017,2020-11.

TE.4.4. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31487

31487
𠄎 30.4 𠄎
VN-F1716

U+31487 above should be unified with U+20BD0 below.

20BD0
𠄎 30.4 𠄎 𠄎 𠄎
UCS2003 GKX-0178.07 T5-2356

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=00383> latest note says:
not unified to U+20BD0 or U+20CBC, irg50.

TE.4.5. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 3158D

3158D
土 32.4 𠄎 𠄎
UTC-02991 SAT-07230

U+3158D above should be unified with U+2BB5B below.

2BB5B
土 32.4 𠄎
GIDC-010

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=00667> latest note says:
not unified to U+2BB5B., IRG 52.

TE.4.6. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 3158F

3158F
土 32.4
𡗗
GDM-00021

U+3158F in previous page should be unified with U+3AAF below.

3AAF
文 67.3
𡗗 𡗘
GHZ-32170.01 T3-263B

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=00668> has no note but mention a similarity with U+575F 坟, not U+3AAF 𡗗. In addition, the lower component in one case is the earth 土 (also used as a radical), while in the other case it is encoded separately as U+30001 𠂇. It does not seem that IRG ever considered unifying U+3158F and U+3AAF.

E.4.7. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31641

31641
女 38.2
𡗗
GXM-00091

U+31641 above should be unified with U+216AB below.

216AB
女 38.2
𡗗 𡗘
UCS2003 TF-2156

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=00875> note says:

Confirmed as simplified form, IRG 53.

not unified to U+216AB or U+216AD, irg50.

E.4.8. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 316B9

316B9
尸 44.6
𡗗
V4-465E

U+316B9 above should be unified with U+663C in next page.

663C
目 72.5
𡗗 𡗘 𡗙 𡗚
G0-5667 T3-2C37 J0-436B K2-3864

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=01024> last note says:

accepted as a separate character based on the Zhuang evidence (#3514), 2020-11.

E.4.9. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31782

31782 𠄎
心 61.4
VN-F0745

U+31782 above should be unified with U+225E9 below

225E9 𠄎 𠄎
心 61.4
UCS2003 TF-267E

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=01473> last note says:

R=61 (心), IRG 52.

not unified to U+225E9, irg50.

E.4.10. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 318BE

318BE 𣎵
木 75.2
GZ-0632404

U+318BE above should be unified with U+673C in next page.

673C 𣎵 𣎵 𣎵 𣎵
木 75.2
G5-4339 HB2-C9F5 T2-2256 K2-3957

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=01666> last note says:

not unified to U+673C, irg50.

E.4.11. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 3196E

3196E 𠄎
水 85.2
T13-222A

U+3196E above should be unified with U+6C5E below.

6C5E 𠄎 𠄎 𠄎 𠄎 𠄎
水 85.3
G0-392F HB1-A845 T1-4A66 J0-5D67 K0-7B71

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=01666> last note says:

not unified to U+6C42, irg51.

In other words, unification with U+6C42 𠄎 was considered, but not U+6C5E which seems to have significantly different components as well.

E.4.12. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31973

31973
水 85.4
沅
UK-10165

U+31973 above should be unified with U+6C85 in next page.

6C85
水 85.4
沅 沅 沅 沅 沅
G0-6364 HB1-A84A T1-4A6B J13-7656 K0-6A36

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=01882> has no note and therefore IRG never considered unifying it with U+6C85 which has a distinct component on the right.

E.4.13. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 3197C

3197C
水 85.5
汪
GXM-00395

U+3197C above should be unified with U+6C6A below.

6C6A
水 85.4
汪 汪 汪 汪 汪
G0-4D74 HB1-A84C T1-4A6D J0-5D6A K0-885C

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=01892> note says:

New evidence, 2020-06.

not unified to U+6C6A or U+2DBFA, irg50.

E.4.14. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 319C8

319C8
水 85.9
溥
T13-2C78

U+319C8 above should be unified with U+6F19 in next page.

6F19
水 85.11
溥 溥 溥 溥
G3-533E HB2-E277 T2-4B3F K1-5C67

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=04991> note says:

Not unified to 溥 U+6EA5, IRG 53.

moved to ws2017, U+6EA5, irg50.

not unified by U+6F19, irg49.

E.4.15. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31B30

31B30 𠄎
火 86.19
T13-374C

U+31B30 above should be unified with U+2A3DB below.

2A3DB 𠄎 𠄎 𠄎
黄 201.10
UCS2003 GKX-1517.16 T4-6A30

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=02429> note says:

new evidence accepted,2020-11.

postponed unified to U+2A3DB, irg51.

E.4.16. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31C4A

31C4A 𠄎
疒 104.3
T13-386B

U+31C4A should be unified with U+75A0 in next page.

75A0 𠄎
疒 104.3
G0-705D

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=02759> has no note and therefore IRG never considered unifying it with U+75A0 which has a distinct component on the lower right. The same page indicates a similarity with U+75C5 病.

E.4.17. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31C4C

31C4C 𠄎
疒 104.4
UK-1046D

U+31C4C above should be unified with U+2DF06 below.

2DF06 𠄎
疒 104.4
GZ-3811101

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=02762> note says:

not unified to U+2DF06, non-cognate, irg51.

E.4.18. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31F37

31F37
舟 137.3
舩
SAT-05571

U+31F37 should be unified with U+26A0C below.

26A0C
舟 137.2
舩 舩 舩
UCS2003 GHZ-53055.10 T6-2D75

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=03627> note says:

not unified to U+2C712 or U+26A0C, irg50.

E.4.19. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31F69

31F69
艸 140.7
葚
GXM-00175

U+31F69 above should be unified with U+26C25 below.

26C25
艸 140.7
葚 葚
UCS2003 GHZ-53205.02

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=03689> note says:

not unified to U+26C25, irg50.

E.4.20. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 31FF5

31FF5
虫 142.9
𧈧
UTC-03113

U+31FF5 above should be unified with U+27310 below.

27310
虫 142.8
𧈧 𧈧 𧈧
UCS2003 GKX-1087.06 T6-6176

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=03875> has no note. This unification hasn't been discussed before, though a glance at the evidence seems to suggest they are non-cognate and thus not unifiable.

E.4.21. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 32112

32112 躉 躉
足 157.17 KC-08160 UTC-03107

U+32112 in previous page should be unified with U+8EA0 below.

8EA0 躉 躉 躉 躉
足 157.17 G3-6F4C HB2-F8C5 T2-7053 K2-644D

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=03825> has no relevant final note related to this, but there is an abundant conversation concerning unification with U+8EA0. It was concluded by Andrew West in the following terms:

NO_UNIFICATION: The two words 躉躉 and 躉躉 or 躉{𠬞薛足} have the same meaning, but U+8EA0 躉 has reading xiè in this context, whereas according to the UTC evidence {𠬞薛足} is a variant of U+8E84 躉 which has reading bì. Therefore I conclude that 躉躉 and 躉躉 or 躉{𠬞薛足} are separate words with same meaning but different pronunciations, and so {𠬞薛足} ≠ 躉.

E.4.21. Clause 34. Code charts ... – CJK Unified Ideographs Ext H – 321CB

321CB 鎚
金 167.6 TC-6635

U+321CB above should be unified with U+2B496 below.

2B496 鎚
金 167.6 JK-65818

Proposed change by Japan:

Please delete from CJK Unified Ideograph Extension H. If the glyph needs to be standardized, please consider registering into IVD.

Not Accepted

See TE.4.1 disposition

For that case, <https://hc.jsecs.org/irg/ws2017/app/?id=04427> note says:
not unified to U+2B496, irg50.

Korea (South): Negative

Comments (GE for General, TE for Technical, ED for Editorial):

TE.1. CJK Unified Ideographs

The charts that should show the following six Hanja chars (horizontally extended) seem missing:

> U+5AB4 kIRG_KSource K6-0004
> U+741F kIRG_KSource K6-0005
> U+74A4 kIRG_KSource K6-0006
> U+74B8 kIRG_KSource K6-0007
> U+7807 kIRG_KSource K6-0008
> U+7A66 kIRG_KSource K6-0009.

Proposed change by Korea:

Add charts showing these 6 chars.

Accepted

These corresponds to horizontal extension and because they involve character additions or glyph changes for existing characters, they don't always get represented in the amendment code charts. They were however documented in the data file CJKSrc.txt part of the amendment. The code charts will be modified to include the pages corresponding to these characters.

USA: Positive with comments

Comments (GE for General, TE for Technical, ED for Editorial):

TE.1. 17.6.2 Standardized variation sequences

Glyphs in the Mathematical Alphanumeric Symbols (and a few in Letterlike Symbols block) are in the roundhand style, but the chancery style is more common today. The two styles were unified when the characters were originally encoded, but they are now not considered interchangeable. We recommend 52 variation sequences be added to distinguish roundhand from chancery style, to bring the 10646 list of standardized variation sequences back into synch with those published in Unicode 14.0. (Reference: L2/20-275 at <https://www.unicode.org/L2/L2020/20275r-math-calligraphic.pdf>).

Proposed change by USA:

Add the following variation sequences to <https://standards.iso.org/iso-iec/10646/ed-6/en/UCSVariants.txt>:

1D49C FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL A
212C FE00; chancery style; # SCRIPT CAPITAL B
1D49E FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL C
1D49F FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL D
2130 FE00; chancery style; # SCRIPT CAPITAL E
2131 FE00; chancery style; # SCRIPT CAPITAL F
1D4A2 FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL G
210B FE00; chancery style; # SCRIPT CAPITAL H
2110 FE00; chancery style; # SCRIPT CAPITAL I
1D4A5 FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL J
1D4A6 FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL K
2112 FE00; chancery style; # SCRIPT CAPITAL L
2133 FE00; chancery style; # SCRIPT CAPITAL M
1D4A9 FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL N
1D4AA FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL O
1D4AB FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL P.
1D4AC FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL Q
211B FE00; chancery style; # SCRIPT CAPITAL R
1D4AE FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL S
1D4AF FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL T
1D4B0 FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL U
1D4B1 FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL V
1D4B2 FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL W
1D4B3 FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL X
1D4B4 FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL Y
1D4B5 FE00; chancery style; # MATHEMATICAL SCRIPT CAPITAL Z
1D49C FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL A
212C FE01; roundhand style; # SCRIPT CAPITAL B
1D49E FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL C
1D49F FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL D
2130 FE01; roundhand style; # SCRIPT CAPITAL E

2131 FE01; roundhand style; # SCRIPT CAPITAL F
1D4A2 FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL G
210B FE01; roundhand style; # SCRIPT CAPITAL H
2110 FE01; roundhand style; # SCRIPT CAPITAL I
1D4A5 FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL J
1D4A6 FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL K
2112 FE01; roundhand style; # SCRIPT CAPITAL L
2133 FE01; roundhand style; # SCRIPT CAPITAL M
1D4A9 FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL N
1D4AA FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL O
1D4AB FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL P
1D4AC FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL Q
211B FE01; roundhand style; # SCRIPT CAPITAL R
1D4AE FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL S
1D4AF FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL T
1D4B0 FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL U
1D4B1 FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL V
1D4B2 FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL W
1D4B3 FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL X
1D4B4 FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL Y
1D4B5 FE01; roundhand style; # MATHEMATICAL SCRIPT CAPITAL Z

Accepted

TE.2. 28 Named sequences

Certain sequences used for units of orthography need to be identified for IDNA and ICANN work on label generation rules. Named sequences will make them identifiable. We recommend 19 named sequences be added for Devanagari, Bengali, Gurmukhi, and Oriya, to bring the 10646 list of named sequences back into synch with those published in Unicode 14.0. (Reference: L2/20-055 at <https://www.unicode.org/L2/L2020/20055-comp-excl-seq.txt> and L2/19-372 at <https://www.unicode.org/L2/L2019/19372-comp-exclusion-seq.pdf>).

Proposed change by USA:

Add the following 19 named sequences to NUSI:

DEVANAGARI SEQUENCE FOR LETTER QA; 0915 093C
DEVANAGARI SEQUENCE FOR LETTER KHHA; 0916 093C
DEVANAGARI SEQUENCE FOR LETTER GHHA; 0917 093C
DEVANAGARI SEQUENCE FOR LETTER ZA; 091C 093C
DEVANAGARI SEQUENCE FOR LETTER DDDHA; 0921 093C
DEVANAGARI SEQUENCE FOR LETTER RHA; 0922 093C
DEVANAGARI SEQUENCE FOR LETTER FA; 092B 093C
DEVANAGARI SEQUENCE FOR LETTER YYA; 092F 093C
BENGALI SEQUENCE FOR LETTER RRA; 09A1 09BC
BENGALI SEQUENCE FOR LETTER RHA; 09A2 09BC
BENGALI SEQUENCE FOR LETTER YYA; 09AF 09BC
GURMUKHI SEQUENCE FOR LETTER LLA; 0A32 0A3C

GURMUKHI SEQUENCE FOR LETTER SHA; 0A38 0A3C
GURMUKHI SEQUENCE FOR LETTER KHHA; 0A16 0A3C
GURMUKHI SEQUENCE FOR LETTER GHHA; 0A17 0A3C
GURMUKHI SEQUENCE FOR LETTER ZA; 0A1C 0A3C
GURMUKHI SEQUENCE FOR LETTER FA; 0A2B 0A3C
ORIYA SEQUENCE FOR LETTER RRA; 0B21 0B3C
ORIYA SEQUENCE FOR LETTER RHA; 0B22 0B3C

Accepted

TE.3. 34.2 Geometric Shapes Extended

The name of U+1F7D9 in the CDAM is NINE POINTED STAR. In order to be consistent with other character names, “white” should be inserted so the character name is “NINE POINTED WHITE STAR.” (Reference: L2/21-144R at <http://www.unicode.org/L2/L2021/21144r-nine-pointed-star.pdf>).

Proposed change by USA:

Change the name to NINE POINTED WHITE STAR.

Accepted

TE.4. 34.2 Arabic Extended-B

In accordance with comment TE.5 below, four Quranic characters currently located in the BMP should be moved to the SMP.

Proposed change by USA:

Move the following three characters into the new block (TE.5 below) as noted:

U+0895 ARABIC SMALL LOW WORD SAKTA to U+10EFD

U+0896 ARABIC SMALL LOW WORD QASR to U+10EFE

U+0897 ARABIC SMALL LOW WORD MADDA to U+10EFF.

Accepted

TE.5. 34.2 Arabic Extended-C (new block)

More proposals for new Arabic characters are expected in the near future. Since only a few spots are available in the BMP, we recommend reserving those spots for modern orthographies, and moving historical and technical characters to the SMP. In order to accommodate such signs, a new Arabic block is needed in the SMP.

Proposed change by USA:

Create a new block, “Arabic Extended-C with the code point range U+10EC0..U+10EFF, and add characters as listed in TE.4 above.

Accepted

TE.6. 34.2 Kawi – removing U+11F3C

Research demonstrates that KAWI VOWEL SIGN VOCALIC L only attested once and its use is not clear (see L2/21-180 at <http://www.unicode.org/L2/L2021/21180-kawi-cognate.pdf> and L2/21-176 at <http://www.unicode.org/L2/L2021/21176-kawi-vowel-sign-vocalic-l.pdf>). We recommend the character be removed but the space be reserved until more information is provided.

Proposed change by USA:

Remove U+11F3C KAWI VOWEL SIGN VOCALIC L.

Accepted

TE.7. 34.2 Kawi -name change for U+11F42

The name CONJOINER is a more appropriate name than SUBJOINER for scripts that have pre- or post-base conjunct forms in addition to subjoined forms.

Proposed change by USA:

Change the name of U+11F42 KAWI SUBJOINER to KAWI CONJOINER.

Accepted

TE.8. 34.2 CJK Unified Ideographs Extension H – source T13-3647

Based on the IRG recommendation WG2 N5165 (L2/21-158) and the CJK editorial group editorial report (IRGN2506 at <https://appsrv.cse.cuhk.edu.hk/~irg/irg/irg57/IRGN2506Confirmed.pdf>), we recommend T13-3647 be removed.

Proposed change by USA:

Remove code point corresponding to T13-3647.

Accepted

See also comment T1 from China. The code point in the amendment is U+31A86.

TE.9. 34.2 CJK Unified Ideographs Extension H – source GZ-4322205

Based on the IRG recommendation WG2 N5165 (L2/21-158) and the CJK editorial group editorial report (IRGN2506 at <https://appsrv.cse.cuhk.edu.hk/~irg/irg/irg57/IRGN2506Confirmed.pdf>), we recommend GZ-4322205 be added.

Proposed change by USA:

Add code point corresponding to GZ-4322205.

Accepted

See also comment T1 from China.

Editorial modifications by project editor

The following editorial modification corresponds to clarification concerning some CJK IRG sources corresponding to CJK Extension H.

G sources based on:

<https://appsrv.cse.cuhk.edu.hk/~irg/irg/irg49/IRGN2228WS2017China2017-09-17.doc>

- GHC HanYu Dacidian (汉语大词典), 1986 年, 上海辞书出版社, 9187.11 / 1988-1993 年, 汉语大词典出版社 (found in 1988), ISBN 7-5432-0001-5/H.2
- GKJ China National Committee for Terms in Sciences and Technologies (CNCTST), names of geology, botanical and zoology terms, WG2 N5072, Chinese medicine characters, IRG N2228.
- GXM Human-name characters. Source from Public Order Administration, The Ministry of Public Security of the People's Republic of China.
- GZA-1 A Vibrant and Unbroken Transmission - Filial Piety and Zhuang Funeral Songs 生生不息的传承·孝与壮族行孝歌之研究, 2010 年北京民族出版社, ISBN 978-7-1051-0648-6
- GZA-2 Annotated Long Zhuang Morality Songs 壮族伦理道德长诗传扬歌译注, 2005 年广西民族出版社, ISBN 7-5363-4922-X
- GZA-3 Compendium of Old Zhuang Folksong Texts - Wooing Songs vol. 1 - Liao Songs 壮族民歌古籍集成·情歌(一) 嘹歌, 1993 广西民族出版社, ISBN 7-5363-2714-5
- GZA-4 Compendium of Old Zhuang Folksong Texts - Wooing Songs vol. 2 - Fwen Nganx 壮族民歌古籍集成·情歌(二) 欢, 1997 年南宁广西民族出版社, ISBN 7-5363-3298-X
- GZA-5 Zhuang Character Database - Internet Collection 方块壮字资料库·网络集 2011 年广西大学, 桂 ICP 备 05000940 号
- GZA-6 Zhuang Proverbs from China 中国壮族谚语 2015 广州世界图书出版广东有限公司, ISBN 978-7-192-0492-1
- GZA-7 Ancient Remembrance - Zhuang Creation Myth Songs 远古的追忆·壮族创世神话古歌研究, 2012 年北京民族出版社, ISBN 978-7-1051-2242-4

V source update:

- VN Vietnamese Nôm characters not defined by V0 to V4 sources, major sources are: Nguyễn Quang Hồng. "Tự Điển Chữ Nôm Dẫn Giải". Hà Nội, 2014, and Anthony Trần Văn Kiệt "Giúp đọc Nôm và Hán Việt", Đà Nẵng, 2004