

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

Doc. Type: Disposition of comments

Title: Disposition of comments on ISO/IEC CD 10646 5th edition

Source: Michel Suignard (project editor)

Project: JTC1.02.10646.00.00.00.06

Status: For review by WG2

Date: 2019-02-07

Distribution: WG2

Reference: SC2 N4635 N4646 N4416 WG2 N5013 N5014 N5016 N5018 IRG N2291

Medium: Paper, PDF file

Comments were received from China, Ireland, Japan, Korea (ROK), UK, and USA. The following document is the disposition of those comments. The disposition is organized per country.

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

Based on the number of negative ballots (5), and the number of conflicting demands, the editor has not tried to create dispositions that would recognize a consensus and allow a fast path to a draft amendment (DAM). Instead, the intent is to have a new CD ballot (CD.2) that will include all non-controversial technical and editorial comments, and many new non-controversial repertoire additions proposed since last WG2 that are supported by fully formed proposals for addition. Based on the disposition, it is likely that only one of the negative ballots could be reversed: Korea (ROK) because all the NB comments were accommodated.

Beyond the ballot comments, another expert contribution, WG2 N5018 written by Peter Constable, was considered in improving the 6th edition. The contribution is focusing on the terms and definition, and the clause 12 related to ISO/IEC identification. The editor felt that incorporating the results of these comments into the next CD version would be productive. As such, these suggestions are disposed in the last part of this document and create modest but useful changes to the text.

Concerning requests to remove repertoires, the dispositions were organized as follows:

- Remove repertoire that were requested to be removed by all parties. This concerns the Latin letters with overcurl: A7D0-A7D9.
- Preserve repertoires that some parties asked to be removed and some other parties asked to be preserved. This concerns CJK ideograph (Gonche) 9FF0-9FF6, Latin letters A7C0-A7C1, Zanabazar Square 11A48-11A49, Counting Rods 1D379-1D37D, Geometric shapes: 1F7E0-1F7EB, Symbols: 1F90D-1F90E, 1F979.

- Preserve the initial CJK Extension G repertoire and reorder (minor) it according to IRG #51. It is expected that IRG will keep working on feedback on CJK ext. G until their meeting #52. This also insures a reasonable stable reference (the reordered parts are extremely limited on scope) on which various IRG experts can keep building feedback.

There are two proposed addition as a direct consequence of these dispositions: 1DFA COMBINING OVERCURL and 1E94B ADLAM NASALIZATION MARK. Details about these changes are provided in the next pages. The other indirect additions (i.e. not related to these dispositions) are also briefly described in the next pages; details can be found in the referenced proposal documents.

In general, for the CD.2 ballot, NBs and liaison organizations are invited to repeat any ballot comments that were not accommodated to their satisfaction either by repeating their ballot text in the new ballot or referencing their comments indicated in these dispositions.

The following items reflect the repertoire changes to the CD between CD.1 and the planned CD.2

Characters added to CD as disposition of comments (supporting document # appended):

1DFA COMBINING OVERCURL (WG2 N4907)
1E94B ADLAM NASALIZATION MARK (WG2 N5022)

Characters removed from CD as disposition of comments:

A7D0 LATIN SMALL LETTER A WITH OVERCURL
A7D1 LATIN SMALL LETTER E WITH OVERCURL
A7D2 LATIN SMALL LETTER I WITH OVERCURL
A7D3 LATIN SMALL LETTER M WITH OVERCURL
A7D4 LATIN SMALL LETTER N WITH OVERCURL
A7D5 LATIN SMALL LETTER R WITH OVERCURL
A7D6 LATIN SMALL LETTER S WITH OVERCURL
A7D6 LATIN SMALL LETTER T WITH OVERCURL
A7D8 LATIN SMALL LETTER U WITH OVERCURL
A7D9 LATIN SMALL LETTER Y WITH OVERCURL

Characters for which removal was requested by some parties but stay (pending further comments in CD.2), one character moved: TROLL from 1F979 to 1F9CC

A7C0 LATIN CAPITAL LETTER THORN WITH DIAGONAL STROKE
A7C1 LATIN SMALL LETTER THORN WITH DIAGONAL STROKE

11A48 ZANABAZAR SQUARE CLUSTER-INITIAL LETTER LA
11A49 ZANABAZAR SQUARE CLUSTER-INITIAL LETTER SA

1D379 SOUTHERN SONG COUNTING ROD UNIT DIGIT FOUR
1D37A SOUTHERN SONG COUNTING ROD UNIT DIGIT FIVE
1D37B SOUTHERN SONG COUNTING ROD UNIT DIGIT NINE
1D37C SOUTHERN SONG COUNTING ROD TENS DIGIT FIVE
1D37D SOUTHERN SONG COUNTING ROD TENS DIGIT NINE

1F7E0 LARGE ORANGE CIRCLE
1F7E1 LARGE YELLOW CIRCLE
1F7E2 LARGE GREEN CIRCLE
1F7E3 LARGE PURPLE CIRCLE
1F7E4 LARGE BROWN CIRCLE
1F7E5 LARGE RED SQUARE
1F7E6 LARGE BLUE SQUARE
1F7E7 LARGE ORANGE SQUARE
1F7E8 LARGE YELLOW SQUARE
1F7E9 LARGE GREEN SQUARE
1F7EA LARGE PURPLE SQUARE

1F7EB LARGE BROWN SQUARE
1F90D WHITE HEART
1F90E BROWN HEART
1F9CC TROLL

30000..31349 CJK UNIFIED IDEOGRAPHS EXTENSION G

Repertoire reordering (minor) through this ballot

30000..31349 CJK UNIFIED IDEOGRAPHS EXTENSION G (see details in comment GE2 from Japan)

CJK Repertoire with K source updates (see details in ROK ballot disposition):

3EAC (delete K source), 8C6C (modified K source, same glyph), 248F2 (new K source, new glyph),
27CEF (new K source, new glyph), 2EB7E (modified K source, same glyph), 2EB89 (modified K source, same glyph)

Characters added to CD as UTC proposed additions, existing blocks (supporting document # appended):

0B55 ORIYA SIGN OVERLINE (WG2 N5023)
A82C SYLOTI NAGRI SIGN ALTERNATE HASANTA (WG2 N5024)
16FF0 VIETNAMESE ALTERNATE READING MARK CA (WG2 N4915, N5011, and N5026)
16FF1 VIETNAMESE ALTERNATE READING MARK NHAY
1F1AD MASK WORK SYMBOL (WG2 N5027)
1F8B0 ARROW POINTING UPWARDS THEN NORTH WEST (WG2 N5028)
1F8B1 ARROW POINTING RIGHTWARDS THEN CURVING SOUTH WEST

Characters added to CD as UTC proposed additions, new blocks (supporting document # appended):

Lisu Supplement 11FB0-11FBF (WG2 N5025)

11FB0 LISU LETTER YHA

Symbols for Legacy Computing 1FB00-1FBFF (WG2 N5028)

1FB00 BLOCK SEXTANT-1
...
1FB92 UPPER HALF INVERSE MEDIUM SHADE AND LOWER HALF BLOCK
1FB94 LEFT HALF INVERSE MEDIUM SHADE AND RIGHT HALF BLOCK
...
1FBCA WHITE UP-POINTING CHEVRON
1FBF0 SEGMENTED DIGIT ZERO
...
1FBF9 SEGMENTED DIGIT NINE

Noted

See comment TE.4 from US.

The character will be balloted again in CD.2. But Ireland needs to make a separate proposal for this character as requested by the US. The location was changed to 1F9CC for type consistency within the block.

T6. Page 1662, Row 1F90: Supplemental Symbols and Pictographs

Ireland affirms its request for the encoding of 1F9AB SQUIRREL. The attempt to unify squirrel with CHIPMUNK is an inappropriate over-extension of squirrel for the source character CHIPMUNK which was encoded for compatibility with the Windows webdings character set.



1F9AB SQUIRREL.

Noted

The case for SQUIRREL is weak as current vendor implementation do not differentiate between the two forms. There is no formal proposal for this character addition.

T7. Page 2671, Row 3050: CJK Unified Ideographs Extension G

Ireland requests the removal of U+3050C (UK-02790) from CJK Unified Ideographs Extension G as it is unifiable with U+22C3A. See attached image. (The T glyph form is anomalous, but as it has the reading fū it should be the same character as UK-02790.).



Not accepted

It is true that U+3050C 損 looks a lot like V2-7729. But the alternative solution may be to move V2-7729 to Ext G at the position U+3050C. This should be discussed by IRG experts at IRG #52 before deciding.

Editorial comments:

E1. Page 295, Row 20A: Currency Symbols

Ireland recommends that the glyph for 20BF BITCOIN SIGN be replaced with a more generic shape. It has always (since the advent of the EURO SIGN) been UCS practice to use glyphs for currency symbols which harmonize with the Times-style glyphs used for Row “000 C0 Controls and Basic Latin”. We can supply a suitable glyph.



Not accepted

While it may be desired to harmonize with a Times-style glyph, the proposed picture may be a less optimal design in many aspects (placement of the bars, size of the bars, etc...). The current glyph is:



E2. Page 1407, Row 1343: Egyptian Hieroglyphs Format Controls

Ireland remains convinced that the glyphs for these characters must be revised. Control characters in the UCS typically have identifying letter abbreviations in their dotted-box glyphs. The exception to this is the Ideographic Description Characters, which are intended to be “visibly displayed graphic characters, not invisible composition controls” as the annotation to the code chart states. The glyphs of the Egyptian characters in this PDAM are hardly distinguishable from those of the Ideographic Description Characters, and we believe this

Japan: Negative

General, Technical, and Editorial comments (noted as GE, TE, or ED)

Technical comments

TE.1. Page 1, 2. Normative references.

In clause 2 “Normative references”, some documents even other than Unicode Technical Reports have the link to www.unicode.org site. It seems that this change has been made following the recommendation from ISO/IEC JTC 1/SC2/WG 2 meeting #67 (Recommendation M67.21.) However, this recommendation is not endorsed by ISO/IEC JTC 1/SC 2. Therefore, Japan requests SC 2 to deliberate this recommendation.

The concerns from Japan about proposed text is there is no guarantees that the documents referred are maintained consistently with this standard, because it may be updated by the organization not under control by this committee. Proposed change by Japan.

Change this clause back to the one in the previous edition.

Also change the URL of documents other than Unicode Technical Reports throughout the documents back to the one starting from “<http://standards.iso.org/iso-iec/10646/...>”

Not accepted

As mentioned by Japanese NB, the change was done according to recommendation M67.21 which was adopted unanimously by WG2 (with Japan experts present). The lack of endorsement by SC2 looks like an oversight and as suggested by Japan, it should be deliberated at the next SC2 meeting.

About concerns with the maintenance consistency for standard, it should be noted that the Unicode standard has a much better maintenance story than ISO. Unicode standard versions have been maintained since version 1.1 published in 1993, and datafiles go back to 1995. In that aspect, Unicode has proven to be an excellent data repository. At the same time, typically ISO only maintains availability of the last version of a standard.

While this aspect can be discussed at the next SC2 plenary in June 2019, it seems advisable to maintain the current status.

Finally, it should be noted that unlike what is suggested in the ‘Propose change by Japan’, ISO/IEC 10646 has historically contained many URLs linked to documents other than Unicode Technical Report. Therefore acting on that requested change would be impossible, because many of these URLs point to resources not available under “<http://standards.iso.org/iso-iec/10646/...>”

GE2. Page 2655, 33.5 Code charts and lists of character names – CJK Extension G

CJK Unified Ideograph Extension-G (“CJK-G” in short) code chart has been updated before SC 2/WG 2 meeting #67 in London following the discussion by IRG meeting #50, then it is adopted in this ballot text.

However, Japan has the concerns with the following.

- Japanese experts reviewed updated CJK-G code chart and found that so many things were not correctly updated following the discussion record from IRG meeting #50. Note Japan expert sent this feedback to IRG convener and technical editor with copying WG2 project editor.
- IRG, then, reviewed above feedback at IRG meeting #51 and issued new report with discussion records. Japanese experts carefully checked this report and confused very much because it concluded that there were so many incomplete or wrong records in the report from IRG meeting #50. On the other hand, many changes has been made on the rules IRG decided before, such as the unification rules, the way of stroke counting and the choice of indexed radicals.

Japan is seriously concerned with such ad hoc operation by IRG. Therefore, Japan still thinks Ext G are not matured yet. Japan would like to suggest IRG re-considering its procedure before proposing new CJK Unified Ideographs Extensions.

By the way, during this CD ballot period, CJK-G was discussed in IRG meeting #51 held in October 22nd thru 26th. See the recommendation IRG M51.2 in SC2 N4641. This obviously violates the following in ISO/IEC Directives, Part 1.

JA.1.1 Discussion during ballot period

... Documents out for ballot at Committee Stage or any later stage shall not be subject to formal discussion at any working level of JTC 1 during the balloting period.

Proposed change by Japan.

Remove CJK-G from this CD

Not accepted

The opinion of Japanese IRG experts was apparently not shared by other IRG experts. It was felt that the result out of IRG meeting #50 was good enough to progress for ballot in the CD as reflected in recommendation M67.18 (unanimous) and endorsed in SC2 resolution M23.04 (although the adoption of M23.04 was not unanimous, nobody objected or abstained on the CJK Extension G part of the resolution).

The review at meeting IRG#51 was extremely limited, resulting only of code point re-ordering due to radical/stroke update. The bulk of CJK Extension G feedback is still due. While it is not clear whether the level of Ext G discussion during IRG #51 rose to the level of 'formal discussion' the editor will make sure the repertoire is not under ballot during the next IRG meeting (#52 in May 2019 Hong Kong, China).

Note also the text in the IRG#51 editorial report

(<http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg51/IRGN2327WS2015EditorialReport.pdf>) as:

Most of the inconsistencies among IRG#50 discussion record (IRGN2291) and IRGN2308RWS2015v6 raised by Japan were confirmed. Their further process were also decided.

...

The following records are all about discussions on Japan's comments (inconsistencies among IRG#50 discussion record IRGN2291 and IRGN2308RWS2015v6). There are 6 stroke counts and 1 glyph to be changed, the corresponding conclusions are marked in red.

Based on this (extracted from the document linked above), the following changes will be made on the CD.2:

3010B 𪛗
𠂇 15.12
GZ-3542301 SC goes to 13, code point moved to U+3010D

3041C 𪛘
𠂇 57.9
GZ-1912202 SC goes to 8, code point moved to U+3041A

309CE 𪛙
石 112.9
GZ-1861301 SC goes to 8, code point moved to U+309C9

30A35 𪛚
禾 115.10
GZ-3271501 SC goes to 11, code point moved to U+30A37

30DCA 𪛛
貝 154.10
GZ-4412301 SC goes to 12, code point moved to U+30DD1

30E21 𪛜
足 157.9
GZ-1231301 glyph change to 𪛜, no R or SC change

30E37 𪛝
足 157.17
GZ-2171101 SC goes to 18, code point unchanged

Resulting code chart is in <http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg51/IRGN2327Codechart.pdf>

IRG members have accepted to conduct further review of CJK Extension G and corresponding updates can be agreed upon at the next IRG meeting (#52) and incorporated into the 6th edition.

TE.3. Page 1076, 33.5 Code charts and lists of character names - Gongche” characters

These “Gongche” characters used for musical notation as the symbol have different characteristic and different usage from CJK unified ideographs. Further, the differences of shape from base characters cannot be considered in the framework of existing unification process. The rule that the character having tiny slash should not be unified with the base character is not consistent with the unification rule adopted so far. Therefore, these characters should not be encoded as CJK Unified Ideograph.

Proposed change by Japan.

Assign these characters into the code point on the block for script and symbols, not CJK ideographs

Not accepted

There have been several opinions on how to encode these characters, either as symbols, CJK characters on their own block, or integrated into an existing CJK block. The 2nd solution was proposed in WG2 recommendation M67.10. This proved impractical for code chart production reason. After all, these characters have radical and stroke count like any other CJK ideograph, information which cannot be easily conveyed in a symbol block. After further feedback, it was considered easier to add them to the more convenient CJK addition area (i.e. end of the URO block starting at U+9FF0).

A further consideration is found in quoting WG2 N4967:

Many characters for Gongche Notation have the same appearances with Chinese ideographs exactly, so the vast majority of them could be used isomorphic ideographs to indicate. Hong Kong SARG once submitted two ideographs which are just used in the lyrics of Yueju Opera as UNCs in IRGN1405R. And UTC has submitted other two ideographs which are used in the lyrics of Kunqu Opera and traditional Gongche Notation to WS2017.

This means that there no intrinsic differences between ideographs used in Gonche Notation that are isomorphic or not. After all, many CJK Unified ideographs are also used in symbolic context (such as digits).

Based on this, it seems preferable to keep the proposed encoding as it is.

TE.4. Page 1849, 33.5 Code charts and lists of character names – U+2278B

The glyph on G column of U+2278B should not be changed, because proposed glyph cannot be considered to be unified with existing one.

Proposed change by Japan.

Do not change the glyph. In case this is error correction, add the information in “Annex P. Additional information on CJK Unified ideographs.”.

Not accepted

For reference, code chart extract in ISO/IEC 5th edition:

2278B
心 61.9 僣 僣
UCS2003 GHZ-10207.12

As proposed in CD 6th edition:

2278B
心 61.9 僣 僣
UCS2003 GHZ-10207.12

While an entry to Annex P may be desired, it does not prevent fixing what is clearly a glyph error for the original source. See WG2 N4988 for further details.

TE.5. 33.5 Code charts and lists of character names – New G sources

As proposed in ISO/IEC JTC 1/SC 2/WG 2 N4988, the source reference in G column of following CJK unified characters are changed.

- from GKX-0631.02 to GHZ-31665.01 for U+3CFD
- from GE-313D to GZFY-28665 for U+6FF9
- from GE-3952 to GLK-421274 for U+809E
- from GE-3D37 to GHZR-63304.09 for U+891D

- from GHZ-10761.12 to GHF-0229 for U+21D4C

The source reference is normative information and this change may cause the compatibility problem on the data transformation table with existing implementation, therefore this change should be avoided.

Proposed change by Japan.

Do not change the source reference of these 5 glyphs on G column.

Not accepted

There are precedents to this kind of changes (including changes affecting J sources), and the requested was made by China (WG2 N4988) which should be well aware of possible impacts on data processing for its own sources. Source information typically determines the identity of encoded characters and may not necessarily be used as primary source for data transformation.

TE.6. Page 2768, A.5.10 390 MOJI-JOHO-KIBAN IDEOGRAPHS-2016 and A.5.11 391 MOJI-JOHO-KIBAN IDEOGRAPHS-2018

In IVD registration of Moji_Joho collection ([Registration of additional sequences in the Moji_Joho collection, dated 2017-12-12](#)), it is found that two different IVSs are registered for one glyph redundantly. As shown in the following figure, [535A E010A] and [2B9E4 E0100] are registered for Moji_Joho MJ059399, also [535A E010B] and [2B9E4 E0101] are registered for Moji_Joho MJ059400.



Proposed change by Japan.

- In JMJKI-2018.txt, delete <535A E010A> and <535A E010B> in the line starting from "535A."
- Associating with this change, add the following statement in Page 2768, Annex A.5.10 390 MOJI-JOHO-KIBAN IDEOGRAPHS-2016 as NOTE, such as "<535A,E010A> and <535A,E010B> have been removed in collection 391."

Accepted

ED.7. Page 34, Table 5 – GHR syntax

Proposed change by Japan.

Change "GHR-dddddd.dd" to "GHR-ddddd.dd."

Accepted in principle

The entry was in fact a duplicate of the previous entry (GHZR) and will be removed entirely along with its description in G sources enumeration.

ED.8. Page 48 Figure 8, Page 49 Fig. 9, Page 51 Fig. 10, Page 52 Fig. 11, Page 53 Fig. 12

Proposed change by Japan.

Change all "unabbreviated names" to "unabbreviated name" in NOTE under these figures.

Accepted

Korea (ROK): Negative

Technical comment:

T1. Page 1997, CJK Unified Ideographs Extension B – U+248F2 and U+3EAC

K6-1002 is currently in U+3EAC. However, it seems proper to move K6-1002 to U+248F2.

If accepted by WG2, KR will send a TTF file containing a modified glyph of K6-1002.

See WG2 N5016 (= IRG N2349) for details.

Proposed change by Korea

Move K6-1002 currently in U+3EAC to U+248F2.

Accepted

The actual reference is K6-1022. The editor needs the font to process this change (another source used in the interim).

Note that document WG2 N5015 is asking to swap the K sources between 2EB7E and 2EB89 because it was an incorrect transcription from the original K submission. This will be accommodated in CD.2.

2EB7E	麥	2EB89	廢
麥 199.9	麥	麻 200.4	廢
	KC-07185		KC-05976

Finally, the K source K1-6B6B was moved to U+27CEF with the glyph that was formerly at U+8C6C. U+8C6C has now a ghost K source KU-08C6C that will need to be changed or formalized.

E1. Page 33, 23.1 List of source references

Insert a blank between X and 1.

X1001, X1002, X1027-1, X1027-2, X1027-3, X1027-4, X1027-5.

Proposed change by Korea

X 1001, X 1002, X 1027-1, X 1027-2, X 1027-3, X 1027-4, X 1027-5.

Accepted

United Kingdom: Negative

General, Technical, and Editorial comments (noted as GE, TE, or ED)

ED.1. Page 12, 6.3.1 Classification (code points)

“The Table 1 summarizes the types”

“The” is unnecessary.

Proposed change by U.K.

Change to “Table 1 summarizes the types”.

Accepted

ED.2 page 24, 16.5 Ideographic description characters

“The Annex I describes them in more details”

“The” is unnecessary.

Proposed change by U.K.

Change to “Annex I describes them in more details”.

Accepted

ED.3. page 24, 16.6.1 General (Variation selectors and variation sequences)

“FVS1 To FVS3” should have lowercase “to”.

Proposed change by U.K.

Change to “FVS1 to FVS3”.

Accepted

TE.4. page 24, 16.6.1 General (Variation selectors and variation sequences)

“A variation sequence whose base character is a CJK Unified ideograph and whose variation selector is from the VARIATION SELECTORS SUPPLEMENT BLOCK is called an ideographic variation sequence.”

UTS #37 (since Revision 10) defines an ideographic variation sequence as being “a sequence of two coded characters, the first being a character with the Ideographic property that is not canonically nor compatibly decomposable, the second being a variation selector character in the range U+E0100 to U+E01EF”. This does not restrict an IVS base character to a CJK unified ideograph, but may be any character with the ideographic property (as long as it is not canonically or compatibly decomposable), including Tangut ideographs.

Proposed change by U.K.

Change to “A variation sequence whose base character is an ideograph which is not canonically or compatibly decomposable and whose variation selector is from the VARIATION SELECTORS SUPPLEMENT BLOCK is called an ideographic variation sequence.”

Add a note that at present the only ideographic variation sequences are for CJK unified ideographs.

Accepted

Though there is an issue with the term ‘ideograph’ which is not fully specified in ISO/IEC 10646. Doing some research, the editor could find some elements in the original proposal in

<https://unicode.org/L2/L2016/16291-ivd-proposal.pdf> but not in the current UTS #37 at

<http://www.unicode.org/reports/tr37/>. The original proposal had an enumeration of the proposed ideographs, UTS #37 specifies an ‘ideographic property’ but without formally establishing the value. Ideally, there should be a reference to an externally defined property (TBD) or a list of those code points.

The Note would read (previous note in sub-clause is renumbered NOTE 1):

NOTE 2 – At present the only specified ideographic variation sequences are for CJK unified ideographs.

TE.5. page 24, 16.6.2 Standardized variation sequences – add link to new Emoji file

“Standardized variation sequences are defined by a machine-readable format that is accessible as a link:

<https://www.unicode.org/wg2/iso10646/edition6/data/UCSVariants.txt>.”

The linked file does not define all standardized variation sequences, but refers to another file named “emoji-variation-sequences.txt” for “Emoji variation sequences”. However there is no link to this file, and it is not in the same directory, so it is not easy for a reader to find this file.

Proposed change by U.K.

Include the emoji-variation-sequences.txt file in the ISO/IEC 10646 data directory, and also include a link to that file in the appropriate subclause.

Accepted

TE.6. page 24, 16.6.2 Standardized variation sequences – restructuring of clause

“The Standardized Variations Sequences contain sequences associated with allowed base characters from the following categories:

Pictographic symbols.”

UCSVariants.txt (=StandardizedVariants-12.0.0.txt) refers to “Emoji variation sequences” which are listed in a separate file (emoji-variation-sequences.txt). Given that the Unicode Standard now defines variation sequences in three different places, as “Standardized variation sequences”, “Emoji variation sequences” and “Ideographic variation sequences”, it would be best for ISO/IEC 10646 to also refer to the three types of variation sequence separately.

Proposed change by U.K.

Remove mention of variation sequences for pictographic symbols from 16.6.2.

Add a new subclause 16.6.3 for Emoji variation sequences, which references emoji-variation-sequences.txt in the data folder for this standard.

Renumber 16.6.3 “Ideographic variation sequences” as 16.6.4.

Update 16.6.1 “General” to reflect the three types of variation sequence defined in the standard.

Accepted

TE.7. page 25, 16.6.3 Ideographic variation sequences – ideograph

“Variations sequences composed of a unified ideograph as the base character ...”

IVS base character is not restricted to a CJK unified ideograph, but may be any character with the ideographic property (as long as it is not canonically or compatibly decomposable), including Tangut ideographs.

Proposed change by U.K.

Change to “Variations sequences composed of an ideograph as the base character ...”.

Accepted

See disposition of comment TE.4 about the definition of ‘ideograph’.

TE.8. page 25, 16.6.3 Ideographic variation sequences – ideographic

“The purpose of IVD registration is to provide a technical solution to represent ideographs variants that are considered unifiable (see Annex S) and should not be encoded in CJK Unified ideographs.”

This only refers to CJK unified ideographs, but IVD registration is open to other ideographic scripts such as Tangut.

Proposed change by U.K.

Change to “The purpose of IVD registration is to provide a technical solution to represent ideographic variants that are considered unifiable (see Annex S for the unification rules for CJK Unified ideographs).”

Note change of ungrammatical “ideographs variants” to “ideographic variants”.

Accepted

See disposition of comment TE.4 about the definition of ‘ideograph’.

ED.9. page 34, 23.2 Source references file for CJK ideographs

“The Table 5 provides the format details”

"The" is unnecessary.

Proposed change by U.K.

Change to "Table 5 provides the format details".

Accepted

ED.10. page 36, 23.3.2 Source reference presentation for CJK UNIFIED IDEOGRAPHS block

"The figure 2"

"The" is unnecessary and "figure" should be capitalized.

Proposed change by U.K.

Change to "Figure 2".

Accepted

ED.11. page 37, 23.3.3 Source reference presentation for CJK UNIFIED IDEOGRAPHS EXTENSION A

"The figure 3"

"The" is unnecessary and "figure" should be capitalized.

Proposed change by U.K.

Change to "Figure 3".

Accepted

ED.12. page 38, 23.3.4 Source reference presentation for CJK UNIFIED IDEOGRAPHS EXTENSION B

"The figure 4"

"The" is unnecessary and "figure" should be capitalized.

Proposed change by U.K.

Change to "Figure 4".

Accepted

ED.13. page 38, 23.3.5 Source reference presentation for CJK UNIFIED IDEOGRAPHS EXTENSION C, D, E, F, and G

"The figure 5"

"The" is unnecessary and "figure" should be capitalized.

Proposed change by U.K.

Change to "Figure 5".

Accepted

ED.14. page 38, 23.4 Source reference presentation for CJK Compatibility ideographs

"The figure 6"

"The" is unnecessary and "figure" should be capitalized.

Proposed change by U.K.

Change to "Figure 6".

Accepted

ED.15. page 40, 24.2 Source reference file for Tangut ideographs

"The Table 6 pro-vides the format details"

"The" is unnecessary and hyphen should be removed.

Proposed change by U.K.

Change to "Table 6 provides the format details".

Accepted

ED.16. page 41, 25.2 Source reference file for Nüshu ideographs

"The Table 7 pro-vides the format details"

Proposed change by U.K.

Remove 1F7E0..1F7EB and 1F90D..1F90E pending further evidence that they are required for encoding.

Not accepted

See comment T3 from Ireland and its disposition.

TE.38. 33.5 Code Charts and lists of character names – CJK Ext G – 30E21

UK submitted UK-01969 to IRG Working Set 2015, but as it was misprinted in the source (*Hànyǔ Dà Zìdiǎn*, 2nd ed., 2010) it was excluded from the CJK Ext. G submission. We have recently provided additional evidence for this character which verifies its glyph form (see http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg51/IRGN2308_UK-01969.pdf).

As this character completes the set of derived simplified characters in *Hànyǔ Dà Zìdiǎn* (2nd ed., 2010), it would be best to include it in Ext. G with the rest of the characters in this set

Proposed change by U.K.

Add UK-01969 (RS 140.9) between 30C5C (T13-3068) and 30C5D (KC-03602).

Not accepted

In <http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg51/IRGN2327WS2015EditorialReport.pdf> (IRG#51 editorial report, it was decided that:

A former WS2015 character (SN03617 UK-01969 𪚥𪚦 项) is accepted for inclusion in WS2017 as an exceptional case. (Also recorded in IRGN2328 Editorial Report on WS2017.)

Typically, IRG is not amenable at reinserting a character that was removed from an extension in an update of such extension. The UK experts are invited to submit this feedback to IRG#52.

TE.39. 33.5 Code Charts and lists of character names – CJK Ext G – 30E21

There is some confusion over the glyph form of 30E21 (GZ-1231301). The evidence shows {𪚥𪚦}, and this was the glyph form shown in the previous version of the Ext. G code chart. This was realised to be a mistake in the evidence, and so the character has been corrected to {𪚥𪚦} in the current code chart. However, some experts have questioned this correction.

Additional evidence from *Nánníng Píng huà Cí diǎn* 南寧平話詞典 [Dictionary of Nanning pinghua dialect] (Jiangsu Jiaoyu Chubanshe, 1997) page 175 confirms that {𪚥𪚦} is the correct glyph form:

【𪚥𪚦】timɿ 突然將腳或手抬起：你踩中我
腳喇，快啲～腳起來 || 壯語有 diem[tiː
m¹]這個詞，古壯字作 𪚥𪚦、點”是抬
(腳)的意思

The RS for 30E21 is given as 157.9, but it should be 157.8 to match the correct glyph form.

Proposed change by U.K.

Do not change the glyph for 30E21.

Correct 30E21 (GZ-1231301) RS to 157.8.

Reorder 30E21 (GZ-1231301) between 30E12 (GHR-73955.06) and 30E13 (UK-02121).

Accepted in principle

However, to avoid further disruption in the Ext G code chart, this change will need to be validated by IRG#52 and if accepted by IRG experts it will be incorporated.

TE.40. 33.5 Code Charts and lists of character names – CJK Ext G – 300C6, 3029A, 30773, and 323C

The following characters have incorrect Radical/Stroke values (see IRGN2269_KR_Resp2R.pdf):

300C6 (KC-00229): RS 9.17 should be 9.16

3029A (KC-00729): RS 32.11 should be 32.12

30773 (KC-02200): RS 85.18 should be 85.19

3123C (KC-04718): RS 196.12 should be 196.13.

Proposed change by U.K.

Correct 300C6 (KC-00229) RS to 9.16. No reordering required for 300C6 (KC-00229).

Correct 3029A (KC-00729) RS to 32.12. Reorder 3029A (KC-00729) between 302A6 (GZ-4921103) and 302A7 (UTC-01219).

Correct 30773 (KC-02200) RS to 85.19. Reorder 30773 (KC-02200) between 30774 (KC-05297) and 30775 (GHR-31928.04).

Correct 3123C (KC-04718) RS to 196.13. Reorder 3123C (KC-04718) between 31240 (GHR-84968.22) and 31241 (GHR-84976.08)..

Not accepted

IRG#51 editorial report at <http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg51/IRGN2327WS2015EditorialReport.pdf> does not address these concerns. The editor has no generic issue with the stroke count and/or the re-ordering. However, all these CJK Ext G matters should be addressed as a whole by IRG.

TE.41. Page 2756, Annex A.1 Collections

Collections 2001 through 2007 are for CJK unified and compatibility ideographs in SIP and TIP (2007 = CJK Unified Ideographs Extension G).

Collection 2000 is for SIP, but there is no collection for TIP.

As CJK Unified Ideographs Extension G continues the 2000 series, we suggest redefining 2000 to cover SIP and TIP..

Proposed change by U.K.

Change Collection 2000 to “SIP and TIP” with code points 20000 through 3FFFD.

Accepted in principle

The exact code point ranges would be 20000-2FFFD and 30000-3FFFD.

ED.42. page 2785, Annex I.3, Table I.1

The IDS examples are not the same as given in the Unicode Standard, which may be confusing for users of the two standards.

Proposed change by U.K.

Change the IDS examples in Table I.1 to match the examples shown in the Unicode Standard 11.0 Figure 18-8.

Accepted

For reference the figure 18-8 is repeated below:

Figure 18-8. Examples of Ideographic Description Characters

U+2FF0	𠄐	U+4EC1	仁	→	𠄐イ二
U+2FF1	𠄑	U+5409	吉	→	𠄑士口
U+2FF2	𠄒	U+8857	街	→	𠄒彳圭子
U+2FF3	𠄓	U+58F9	壹	→	𠄓士一豆
U+2FF4	𠄔	U+56DE	回	→	𠄔口口
U+2FF5	𠄕	U+51F0	凰	→	𠄕几皇
U+2FF6	𠄖	U+51F6	凶	→	𠄖口乂
U+2FF7	𠄗	U+5321	匡	→	𠄗仁王
U+2FF8	𠄘	U+4EC4	仄	→	𠄘厂人
U+2FF9	𠄙	U+5F0F	式	→	𠄙弋工
U+2FFA	𠄚	U+8D85	超	→	𠄚走召
U+2FFB	𠄛	U+5DEB	巫	→	𠄛工从