

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

Doc. Type: Disposition of comments

Title: Draft Disposition of comments on PDAM2 to ISO/IEC 10646 5th edition

Source: Michel Suignard (project editor)

Project: JTC1 02.10646.00.02.00.05

Status: For review by WG2

Date: 2018-03-13

Distribution: WG2

Reference: SC2 N4585 WG2

Medium: Paper, PDF file

Comments were received from the following members: Canada, Ireland, Japan, 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 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 comments of two of the National Bodies that had casted negative votes (Japan and US) have been accommodated (taking into account specific instructions by the NBs on what could reverse their vote); and only one character was added. Therefore, it can be considered that only one negative vote remains (Ireland). Therefore, it seems possible to re-issue a new Pdam ballot creating enough consensus for an enquiry ballot.

To maximize the chance of having the next PDAM2 ballot (Pdam 2.3) reaching large consensus, the intent of this disposition is to remove all controversial repertoire from Amendment 2 and postpone all the removed repertoire, moving it to a new 'bucket'. It is assumed that a successful Pdam2.3 can lead its content to be part of an enquiry ballot (DAM2) without additional discussion. In the same spirit, additions are limited to the strict essential. The DAM2 process could start soon after the SC2/WG2 meeting in June 2018.

The repertoire of the 'bucket' removed from Amendment 2 can be discussed further with the following possible outcomes.

- 1) Characters or blocks reaching large consensus before DAM2 is issued that can be re-integrated into the DAM2 repertoire.
- 2) Characters or blocks that need further technical discussion and can be added to the next vehicle for repertoire addition (CD of 10646 6th edition)
- 3) Characters or blocks that don't get significant support and should just not be included in any CD or PDAM.

The following items reflect the changes to Amendment 2 between pdam2.2 and the planned pdam2.3

Character added:

119E4 NANDINAGARI VOWEL SIGN PRISHTHAMATRA E

Block moved

NANDINAGARI block moved from U+11B90..11BEF to U+119A0..119FF

Characters/Blocks removed from Amendment 2 and postponed for further consideration:

A7C0	LATIN CAPITAL LETTER THORN WITH DIAGONAL STROKE
A7C1	LATIN SMALL LETTER THORN WITH DIAGONAL STROKE
16FF0	VIETNAMESE ALTERNATE READING MARK CA
16FF1	VIETNAMESE ALTERNATE READING MARK NHAY
18B00..18CD5	KHITAN SMALL SCRIPTS (block)
1B300..1B4E5	SHISHU LOGOGRAMS (block)
1B500..1B52E	SHUISHU RADICALS (block)
1E2EC	WANCHO TONE TUP
1E2ED	WANCHO TONE TUP MANG
1E2EE	WANCHO TONE OKOI
1E2EF	WANCHO TONE OKOI MANG
1F93F	BILLIARDS GAME
1F97B	TROLL
1F9A3	MAMMOTH
1F9A4	DODO
1F9A5	SQUIRREL
30000..31389	CJK UNIFIED IDEOGRAPHS EXTENSION G (block)

Names changes

UCS	Old name	New name
2E4F	DOUBLE OBLIQUE HYPHEN WITH FALLING DOTS	MEDIEVAL CORNISH VERSE SEPARATOR
A7BC	LATIN CAPITAL LETTER EGYPTOLOGICAL YOD	LATIN CAPITAL LETTER GLOTTAL I
A7BD	LATIN SMALL LETTER EGYPTOLOGICAL YOD	LATIN SMALL LETTER GLOTTAL I
1F9B9	SUPERVILAIN	SUPERVILLAIN
1F9F1	BRICKS	BRICK
1F9FB	ROLL OF TOILET PAPER	ROLL OF PAPER

Glyph changes through this ballot

A7C2 LATIN CAPITAL LETTER ANGLICANA W

A7C3 LATIN SMALL LETTER ANGLICANA W
13432 EGYPTIAN HIEROGLYPH START AT TOP
13433 EGYPTIAN HIEROGLYPH START AT BOTTOM
13434 EGYPTIAN HIEROGLYPH END AT TOP
13435 EGYPTIAN HIEROGLYPH END AT BOTTOM
1F9B8 SUPERHERO
1F9B9 SUPERVILLAIN
1F9F1 BRICK (following name change)

Glyph changes requested by the Emoji Subcommittee (minor)

1F94F FLYING DISC (change picture inside disc)
1F99B HIPPOPOTAMUS (full body)
1F9EB PETRI DISH (overlapping dish and cover)
1F9ED COMPASS (different needle position)
1F9F0 TOOLBOX (different design, two locks)
1F9F2 MAGNET (different orientation)
1F9F7 SAFETY PIN (closed)

Annotations, cross reference, source reference added/modified

2BD2 GROUP MARK (xref to 2E4B)
AB67 LATIN SMALL LETTER TS DIGRAPH WITH RETROFLEX HOOK (xref to 0282)
9FEB CJK UNIFIED IDEOGRAPH-9FEB (new RS value: 84.13)
9FEF CJK UNIFIED IDEOGRAPH-9FEF (new J source: JMJ-060040)
1F43F CHIPMUNK
1F9A0 MICROBE
1F9E7 RED GIFT ENVELOPE
1FA67 XIANGQI BLACK GENERAL

Canada: Positive with comments

Comments:

GE.1. Ballot question Q.1

It is OK to circulate as a CD. (not a DIS), since the text to be circulated is for a PDAM (not a DAM).

Proposed change by Canada

None.

Noted

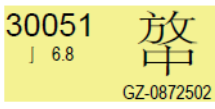
This could be clarified, because while the text is circulated as PDAM (or CD), the next phase is to go to a DAM (or DIS), so the question can be interpreted as requesting whether the PDAM is mature enough to progress to enquiry phase. The editor will consult the SC2 secretariat for clarification.

China: Positive with comments

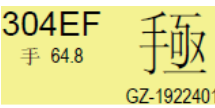
Technical comments:

T.1. CJK Unified Ideographs Extension G

Some characters show wrong radicals or stroke counts, they should be corrected according to technical conclusions made by the IRG#50. Some of them are shown below:

U+30051  6.8 should be changed to 066.0 支.

U+30226  30.12 should be 30.11.

U+304EF  64.8.

Proposed change by China

Apply changes.

Noted

See also comment TE.1 from Japan, TE.2, TE.3, and ED.5 from UK and T15 from USA and their dispositions.

These comments (there seems to be a typo in the comment about U+304EF) should be considered by IRG 50 and result in a new version of CJK Extension which may or may not be brought back into a future DAM2 based on discussion by experts and SC2 members at the next SC2 and WG2 meeting in London (June 2018).

Editorial comment:

E.1. Chess symbols

1FA67 XIANGQI BLACK GENERAL

“= hēi jiāng” should be “= hēi jiàng”..

Proposed change by China

Change annotation.

Accepted

See also comment ED.4 from UK and its disposition

Note this is an annotation for a character not under ballot (1FA67 is part of DAM1).

Ireland: Negative

Ireland disapproves the draft with the technical and editorial comments given below. Acceptance of these comments and appropriate changes to the text will change our vote to approval.

Technical comments:

T1. Page 26, Row 2E0: Supplemental Punctuation

Ireland affirms its strong support for the encoding of 2E4F DOUBLE OBLIQUE HYPHEN WITH FALLING DOTS. This character is found in a range of manuscripts from the 12th to 16th centuries and is a part of the unique Cornish scribal tradition.

Noted

See comment T5 from US and its disposition.

The proposed new name is MEDIEVAL CORNISH VERSE SEPARATOR.

T2. Page 8, Row A72: Latin Extended-D

Ireland affirms its strong support for the encoding of A7C0 LATIN CAPITAL LETTER THORN WITH DIAGONAL STROKE and ACC1 LATIN SMALL LETTER THORN WITH DIAGONAL STROKE. Ireland notes that other Latin letters, such as 0244 Ꝁ LATIN CAPITAL LETTER U WITH STROKE and 0289 ꝁ LATIN SMALL LETTER U BAR alongside A7B8 Ū LATIN CAPITAL LETTER U WITH DIAGONAL STROKE and A7B9 ŭ LATIN SMALL LETTER U WITH DIAGONAL STROKE, are similarly distinguished for use in various languages. The Nordacist tradition is properly served with A7C4 LATIN CAPITAL LETTER THORN WITH STROKE and A7C5 LATIN SMALL LETTER THORN WITH STROKE, but the centuries-old Anglicist tradition (both in manuscript and set in type) would remain unsupported without these two characters. The Anglicist letters are not glyph variants of the Nordacist letters.

Ꝁ ≠ Ū
ꝁ ≠ ŭ

Noted

See comment T1 from US and its disposition.

T3. Page 30, Row A72: Latin Extended-D

Ireland objects to the naming convention for six of the characters on the PDAM. These are as follows:

A7BA LATIN CAPITAL LETTER GLOTTAL A

A7BB LATIN SMALL LETTER GLOTTAL A

A7BC LATIN CAPITAL LETTER EGYPTOLOGICAL YOD

A7BD LATIN SMALL LETTER EGYPTOLOGICAL YOD

A7BE LATIN CAPITAL LETTER GLOTTAL U

A7BF LATIN SMALL LETTER GLOTTAL U

ᶖ ᶗ ᶘ ᶙ ᶚ ᶛ

All six of these letters use the same diacritical mark, which was indisputably devised by Lepsius from the Greek diacritic SPIRITUS LENIS. To name these six characters differently is to ignore this history, and to introduce inconsistency and confusion into the standard. The use of the term “glottal” begs a number of questions as to the phonetic value of the letters (unlike the use of this term in some Ethiopic character names) but consistency

is more important. The best and most accurate names are those initially proposed and accepted as given in N4792R (L2/17-076R2).

A7BA LATIN CAPITAL LETTER A WITH SPIRITUS LENIS

A7BB LATIN SMALL LETTER A WITH SPIRITUS LENIS

A7BC LATIN CAPITAL LETTER I WITH SPIRITUS LENIS

A7BD LATIN SMALL LETTER I WITH SPIRITUS LENIS

A7BE LATIN CAPITAL LETTER U WITH SPIRITUS LENIS

A7BF LATIN SMALL LETTER U WITH SPIRITUS LENIS

If it is impossible to accept such names for a technical reason we would like to know that specific technical reason. Failing that however there are two options, both less satisfactory:

A7BA LATIN CAPITAL LETTER GLOTTAL A

A7BB LATIN SMALL LETTER GLOTTAL A

A7BC LATIN CAPITAL LETTER GLOTTAL I

A7BD LATIN SMALL LETTER GLOTTAL I

A7BE LATIN CAPITAL LETTER GLOTTAL U

A7BF LATIN SMALL LETTER GLOTTAL U

A7BA LATIN CAPITAL LETTER UGARITICIST ALEF

A7BB LATIN SMALL LETTER UGARITICIST ALEF

A7BC LATIN CAPITAL LETTER UGARITICIST YOD

A7BD LATIN SMALL LETTER UGARITICIST YOD

- also used in Egyptology

A7BE LATIN CAPITAL LETTER UGARITICIST WAW

A7BF LATIN SMALL LETTER UGARITICIST WAW

Note that in Ugaritic studies 𐎗 is transliterated as ʾa or â, 𐎙 is transliterated as ʾi or î, and 𐎛 is transliterated as ʾu or û.

Propose Accepted in principle

The proposed names follow the first option, resulting in only two actual name changes and one annotation.:

A7BC 𐎙 LATIN CAPITAL LETTER GLOTTAL I

A7BD 𐎛 LATIN SMALL LETTER GLOTTAL I

- also used in Egyptology as yod

Although the editor thought that the proposed name for the ‘EGYPTOLOGICAL YOD’ reflected better the main intended use.

T4. Page 53, Row 16FF: Ideographic Symbols and Punctuation

Ireland understands that there are many Vietnamese Chữ Nôm ideographs contain these two components within the ideographic square. The proposal document itself states:

However, a rather large number (64) of glyphs, such as U+2A76A [𡗖] composed with the mark are already in Unicode. Compatibility decomposition mappings for these can be provided by the VNPF.

This sort of composition does not seem to be in accord with the principles of ideographic encoding. The two characters on the PDAM may be either components or radicals, but the appropriateness of dynamically composing CJK characters in the way proposed here has not been adequately demonstrated. Nor has the wisdom of using compatibility decompositions. We would propose that a listing of characters with these components be drawn up, and the characters in that list simply be encoded in the normal way. In any case we

do not consider that these characters are good candidates for encoding, and we recommend their removal from ballot pending further research and consensus.

Propose Accepted

See also comment TE1 from UK.

It is true that using compatibility decomposition mapping in the context of CJK Ideographs seem unwise. At the same time, the use of these alternate reading marks could be productive and lead to the encoding of an unnecessary large set if these symbols could be instead used. However, the terminology needs to be improved and it should be made clear that no compatibility decomposition (in the context of ISO/IEC 10646 and Unicode) should be pursued in this context.

These 2 characters (proposed at 16FF0 and 166F1) will be removed from Amendment 2.

T5. Page 83, Row 18B0: Khitan Small Script

Ireland affirms its support for the encoding of the Khitan Small Script as presented here in this ballot. Work to explore the appropriate mechanism for cluster formation is on-going, but since many Khitan specialists cite Khitan Small Script characters in isolation, the set as presented is suitable for encoding.

Noted

See comment GE.1 from UK and comment T6 from US and its disposition.

An initial mechanism for cluster formation was originally part of the proposal and was removed because of controversy. It would be beneficial progress on that aspect before encoding any part of the script.

T6. Page 93, Row 1B30: Shuishu Logograms

With regard to ISO/IEC JTC1/SC2/WG2 N4925 (L2/18-005) "Naming conventions for larger character sets"

Ireland affirms its strong support for the character names for Shuishu logograms as presented in the PDAM. We would oppose any proposal to change these to algorithmically-derived character names.

Noted

See comment GE.2 from UK and comments TE.2 from Japan and T4 from USA and their disposition.

An important part of the Japanese comment relates to the lack of stability for the names of the Shuishu which are classified in N4925 in the category of '1. Meaningful names'. Unlike the other 2 categories described in the same document (2. Catalog names, and 3. Algorithmic names), meaningful names inherently suffer from lack of consensus and therefore tend to be instable. Per ISO/IEC 10646 and Unicode name policy, character names are immutable once encoded. 'Correcting names' requires the use of 'formal aliases' which is cumbersome at best. In many instances it is easier to use Algorithmic names for the formal character names and document the 'meaningful names' in an ancillary file. This was done in the past for Nüshu.

This deserves additional discussion and therefore the encoding of Shuishu can be postponed.

T7. Page 103, Row 1E2C: Wancho

Ireland affirms its support for the four combining characters encoded at 1E2EC..1E2CF. We note that as of this date the UTC Pipeline has not been updated to indicate UTC approval for these characters, which were devised by the script's inventor during the encoding and are currently being used to prepare lexicographical work for new publications.

Noted

See comment T13 from USA and its disposition.

The UTC Pipeline not including these 4 combining characters was based on the Unicode Script ad-hoc report (L2/17-255: <http://www.unicode.org/L2/L2017/17255-script-ad-hoc.pdf>) which stated:

We reviewed this revised proposal, which now includes four tone marks and two short vowels (U+1E2C1 WANCHO LETTER A and U+112E4 WANCHO LETTER ANG).

The creator of the script is a speaker of Northern Wancho, but there are other varieties of Wancho, including Southern Wancho, where the tones vary from Northern Wancho. As a result, the proposed tones

do not mark absolute tone. (Note: In the earlier version of the proposal, tones were marked by doubling or tripling letters.)

The current repertoire – minus the four tones and the two new short vowels – is well attested and has clear evidence of use and could be approved. In our opinion, holes could be left for the 6 characters that have no published examples. Once printed examples are provided, then they can be proposed.

After further discussion, UTC accepted the two short vowels, but required further evidence for the four tone characters.

T8. Page 115, Row 1F90: Supplemental Symbols and Pictographs

Ireland affirms its support for the encoding of 1F93F BILLIARD GAMES, since every vendor has changed the glyph of 1F3B1 BILLIARDS to that of the divination toy EIGHT BALL. It would be wrongly inconsistent for the black and white glyph for that character to differ from the form used by vendors for emoji purposes.

Noted

See comment TE.7 from USA and its disposition.

Not every vendor has changed its glyph. According to <http://unicode.org/emoji/charts/full-emoji-list.html> Samsung still uses the symbol currently proposed at 1F93F for 1F3B1. Therefore, it may be premature to encode another character to disunify a symbol which is currently ambiguous.

T9. Page 115, Row 1F90: Supplemental Symbols and Pictographs

Ireland believes that the glyph of an animal face and the glyph of an animal differ considerably, and that standardizers and vendors should respect this. Accordingly, we believe that either the name of 1F99B be changed to HIPPOPOTAMUS FACE and 1F99D be changed to RACCOON FACE, or that the glyphs be changed to full-body animals for these two characters. An alternative which we would find acceptable would be the encoding of both HIPPOPOTAMUS and RACCOON alongside HIPPOPOTAMUS FACE and RACCOON FACE.

Propose accepted in principle

Besides some animal that are encoded in both face and full body (like pig, rabbit, cow, etc.) many animals can be represented by either a face or the body (like chicken, rooster, bird, etc.). In these cases, it seems better to modify these 2 glyphs to show more than the face of the animal.

The Emoji Sub-Committee has already requested to change the Hippopotamus glyph to show the full-body animal.

T10. Page 115, Row 1F90: Supplemental Symbols and Pictographs

Ireland affirms its strong support for the encoding of 1F98B TROLL.

Noted

See comment TE.10 from USA and its disposition.

T11. Page 115, Row 1F90: Supplemental Symbols and Pictographs

Ireland affirms its strong support for the encoding of 1F9A1 BADGER, 1F9A2 SWAN, 1F9A3 MAMMOTH, 1F9A4 DODO, and 1F9A5 SQUIRREL.

Noted

See comment TE.8 (concerning MAMMOTH and DODO) and TE.9 (SQUIRREL) from USA and their disposition.

There is no controversy concerning the BADGER and the SWAN.

T12. Page 115, Row 1F90: Supplemental Symbols and Pictographs

Ireland maintains its objection to the encoding of the meaningless character 1F9A0 MICROBE. Ireland requests the encoding of the new characters VIRUS and BACTERIA to address the nonscientific and erroneous encoding of the “microbe” invented by the Emoji Subcommittee in order to sell telephones.

Propose not accepted

A microbe is typically defined as a microorganism and can be ambiguously be a bacterium or a virus. (Oxford Dictionary: A microorganism, especially a bacterium causing disease or fermentation). The proposer wanted to keep the ambiguity by using that term. There is nothing inherently incorrect in that proposal. The disparaging comment about the Emoji Subcommittee is not helping, The original proposal is located at: <http://www.unicode.org/L2/L2017/17113-science-emoji.pdf>

An annotation can be added as follows to cover the ambiguity:

** microorganism, intended to cover bacteria, viruses, amoebas, etc.*

T13. Page 115, Row 1F90: Supplemental Symbols and Pictographs

Ireland requests the correction of the name of 1F9B9 SUPERVILAIN to SUPERVILLAIN.

Accepted

This was a typo.

T14. Page 115, Row 1F90: Supplemental Symbols and Pictographs

Ireland objects to the name change of 1F9E7 to RED GIFT ENVELOPE. The red envelope used in East and Southeast Asia is conventionally called RED ENVELOPE, and the informative annotation already explains that it is used as a gift (though this is misspelt “gif” in the PDAM).

Propose not Accepted

There is nothing wrong with the proposed name. The typo in the annotation will be fixed.

T15. Page 115, Row 1F90: Supplemental Symbols and Pictographs

Ireland believes that the encoding of the invented character 1F9C2 SALT SHAKER creates a gap in the UCS and requests a new character PEPPER MILL be added to correct this.

Propose not Accepted

New character request needs a separate contribution.

T16. PDAM 2.2

Ireland objects to the UTC’s publication of the pictographic characters on this PDAM while those characters are under ballot. This effectively invalidates the vote of every National Body participating in SC2 with regard to the technical content of ISO/IEC 10646. SC2 has not been asked, and has not agreed, to rubber-stamp whatever characters the UTC’s Emoji Subcommittee decides to invent in order to help vendors sell their telephones and other devices. That Subcommittee keeps making blunders and mistakes, creating gaps in the set of pictographic characters encoded in the UCS. Oversight is required. We have suggested many times that a fair price for the UTC to pay for the dumping of 50 or 60 characters a year into the UCS for the purposes of emoji would be to allow the addition of a smaller set of characters, like dinosaurs, pepper mills, animal faces, and so on, to make the set of UCS pictographs generally useful for everyone, not just those exchanging colourful emoji characters in text messages and the like. The Emoji Subcommittee may be satisfied with using Google frequency counts and other heuristics as a “scientific” tool to predict what symbols might be popular as emojis, but in effect they are simply using a magic 8-ball to do it. Vendors do not own the UCS. National Bodies have their own expertise, and the UTC is effectively preventing National Bodies from having a say in the population of the UCS with symbols and pictographs.

Propose Out of scope

Ballot comment may be not the optimal place to make such comment.

Editorial comments:

E1. Page 25, Row 2B0: Miscellaneous Symbols and Arrows

Page 25, Row 2B0: Miscellaneous Symbols and Arrows. Ireland suggests adding a cross-reference at 2BD2 GROUP MARK to 2E4B TRIPLE DAGGER.

Accepted

E2. Page 26, Row 2E0: Supplemental Punctuation

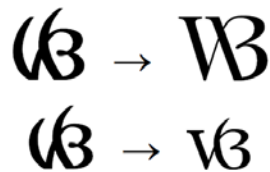
Ireland recommends the addition of an informative note to 2E4F DOUBLE OBLIQUE HYPHEN WITH FALLING DOTS: “used as a verse separator in medieval Cornish manuscripts”.

Propose accepted in principle

See USA comment T5. Based on that comment, the name for the character would be changed to ‘MEDIEVAL CORNISH VERSE SEPARATOR’ which removes the need for additional explanation in annotation.

E3. Page 30, Row A72: Latin Extended-D

After consultation with expert typographers about the appropriate design of the Anglicana W, Ireland requests the following glyph changes for A7C2 and A7C3 respectively. The Anglicana W itself has a number of glyph variants in the manuscript tradition, and the typographic forms shown on the right below are both acceptable within that tradition and offer a modern look which harmonizes well in text.



Propose accepted

Based on receiving an updated font from Ireland.

E4. Page 35, Row AB3: Latin Extended-E

Because the character name of this ligated letter differs somewhat from that of one of its elements, Ireland recommends the addition of a cross reference for AB67 LATIN SMALL LETTER TS DIGRAPH WITH RETROFLEX HOOK to 0282 LATIN SMALL LETTER S WITH HOOK.

Propose accepted

E5. Page 49, 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 could be confusing to users and implementers. After consultation with expert Andrew Glass, we propose that the glyphs below be used, as they conform well to the general principles used in the UCS for control characters.

EVJ ⋮	13430 EGYPTIAN HIEROGLYPH VERTICAL JOINER
EHJ ✱	13431 EGYPTIAN HIEROGLYPH HORIZONTAL JOINER
EST ⌞	13432 EGYPTIAN HIEROGLYPH START AT TOP
ESB ⌟	13433 EGYPTIAN HIEROGLYPH START AT BOTTOM
EET ⌞	13434 EGYPTIAN HIEROGLYPH END AT TOP
EEB ⌟	13435 EGYPTIAN HIEROGLYPH END AT BOTTOM
EOM ⊕	13436 EGYPTIAN HIEROGLYPH OVERLAY MIDDLE
EBS (	13437 EGYPTIAN HIEROGLYPH BEGIN SEGMENT
EES)	13438 EGYPTIAN HIEROGLYPH END SEGMENT

Propose partially accepted

See also comment TE.3 from US.

There is no requirement to have letter abbreviations in glyphs used for control characters, in fact many do not, and the only typical requirement is to use a dotted rectangular box. However, the issue of confusion with Ideographic Description Characters is valid. The following list shows the newly proposed glyphs:

13430	⋮	EGYPTIAN HIEROGLYPH VERTICAL JOINER
13431	✱	EGYPTIAN HIEROGLYPH HORIZONTAL JOINER
13432	⌞	EGYPTIAN HIEROGLYPH START AT TOP
13433	⌟	EGYPTIAN HIEROGLYPH START AT BOTTOM
13434	⌞	EGYPTIAN HIEROGLYPH END AT TOP
13435	⌟	EGYPTIAN HIEROGLYPH END AT BOTTOM
13436	⊕	EGYPTIAN HIEROGLYPH OVERLAY MIDDLE
13437	(	EGYPTIAN HIEROGLYPH BEGIN SEGMENT
13438	)	EGYPTIAN HIEROGLYPH END SEGMENT

E6. Page 115, Row 1F90: Supplemental Symbols and Pictographs

Ireland believes that the glyphs for 1F9B8 SUPERHERO and 1F9B9 SUPERVILLAIN should be replaced with more anthropomorphic figures. We will provide such glyphs in due course, though they are not available at present.

Propose accepted in principle

Based on receiving a font with the updated glyphs.

E7. Page 115, Row 1F90: Supplemental Symbols and Pictographs

Ireland believes that character 1F9F1 BRICKS should be changed to BRICK WALL with the reference glyph retained as on the PDAM. Ireland also requests the addition of a new character BRICK with the glyph of a single brick in perspective. We will provide this glyph in due course.

Propose partially accepted

See also comment TE.11 from USA and its disposition.

BRICKS was changed to BRICK, and the glyph will be updated accordingly.

Japan: Negative

Technical comment:

TE.1. Page 122-1848, sub-clause 33.5 Code charts and lists of character Names – CJK Unified Ideographs Extension G

Japan NB thinks CJK Unified Ideographs Extension G does not yet have the enough quality to be included into PDAM text. There are many open issues in CJK G as described in the document [SC2 N4584] Japan NB separately sent. IRG still needs to work on these issues in order to improve the quality.

According to IRG recommendation M49.2 from last IRG meeting, IRG still continue for one more round of review for WS (Working Set) 2015, after it has been submitted as CJK G to WG2 for the inclusion into PDAM. Japan NB thinks this is incorrect procedure. It should be submitted after the completion of development work within IRG.

Proposed change by Japan:

Drop CHK G.

Proposed Accepted in principle

See also comment T1 from China, TE 2, TE3, ED5 from UK, and TE.15 from USA.

The editor does not think it was an incorrect procedure to include CJK Extension G in pdam2.2. A PDAM ballot is a committee level process and having the CJK Extension G content in ballot allows to widen the scope of balloting members beyond the IRG members. Furthermore, it does not prevent IRG members to keep providing constructive feedback. It is also clear that many IRG members wanted closure on the Working Set 2015 on which CJK Extension G is based to start spending more time on the next working set: WS2017. Putting Ext G on the ballot was a way to expedite the feedback process.

However, it is also essential to get a chance to bring closure to Amendment 2. Therefore, it is acceptable to remove CJK Extension G from a future pdam2.3 to achieve better chance of a unanimous (or as close as possible) position to allow a DAM2 to emerge as a committee consensus.

TE.2. Page 91-98, sub-clause 33.5 – Shuishu Logograms

The phonetic value is used in the character name of Shuishu character, e.g., "QYED7-A" in the character name of U+1B300 "SHUISHU LOGOGRAM QYED7-A." However, Japan NB thinks the phonetic value is not appropriate as the character name of Shuishu character with the following reasons.

- The phonetic value is not essential information to identify the Shuishu character.
- Since Shui language has several dialects, it is difficult to choose a single immutable phonetic value that all Shuishu users are satisfied.

The romanization used for the phonetic value of Shuishu language is not yet stabilized.

Proposed change by Japan:

Remove the phonetic value in the character name. We propose to have further discussion to decide the character name. Or, one of alternatives is to simply use the UCS code point like "SHUISHU LOGOGRAM-1B300."

Proposed Accepted in principle

See comment T6 from Ireland TE.4 from USA and their disposition.

Indeed, further discussion would be useful.

TE.2. Page 29, sub-clause 33.5 – CJK Unified Ideographs

The source reference for U+9FEF is incorrect.

Proposed change by Japan:

Change JMJ-060049 to JMJ-060040.

Proposed Accepted

UK: Positive with comments

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

ED.1. Clause 24 Source references for Tangut ideographs

UTN42 is missing from the list of Tangut source references and Table 6.

Proposed change by UK:

Add "UTN42" to the list of Tangut source references, and add the format for "UTN42" references to Table 6.

Accepted in principle

This was already done in Amendment 1, so there should be no need to update ISO/IEC 10646.

TE.1. Clause 33 Code charts and list of character names – Ideographic Symbols and Punctuation (Vietnamese reading marks)

16FF0 VIETNAMESE ALTERNATE READING MARK CA

16FF1 VIETNAMESE ALTERNATE READING MARK NHAY

There is no consensus among IRG members to encode these two proposed characters. A number of Chinese sources that we have checked do not consider these to be diacritical marks appended to CJKV ideographs but integral components of CJKV ideographs representing particular ideographic composition strategies, both in Vietnam and in China. Both these "marks" are components in many already-encoded CJKV ideographs, and encoding these characters will therefore result in multiple encodings for these characters, which will cause confusion among end users and problems with data interchange and processing.

Proposed change by UK:

Remove 16FF0 and 16FF1 from the amendment pending further study by experts as to whether they should be encoded as separate characters or not.

Propose Accepted

See also comment T4 from Ireland.

ED.2. Clause 33 – CJK Unified Ideographs

U+9FEB: stroke count should be 84.13 according to IRG stroke counting rules.

Proposed change by UK:

Change stroke count of U+9FEB to 84.13.

Accepted

GE.1. Clause 33 – Khitan Small Script

We affirm our support for keeping the current repertoire for Khitan Small Script in Amd. 2. As discussed at WG2 Meeting 66 in Hohhot (see WG2 N4905 "Khitan Ad Hoc Report (Hohhot, Inner Mongolia)") most users do not need to display Khitan Small Script clusters, but only require use of individual Khitan Small Script glyphs in isolation or presented in linear format. Therefore it is useful to encode the repertoire of glyphs first, without any format characters, and leave the cluster-forming mechanism for future study and agreement by experts.

Proposed change by UK:

None.

Noted

See also comment T5 from Ireland and TE6 from USA and its disposition.

An initial mechanism for cluster formation was originally part of the proposal and was removed because of controversy. It would be beneficial to see progress on that aspect before encoding any part of the script.

GE.2. Clause 33 – Shuishu Logograms

We affirm our support for the names used for Shuishu Logograms. As discussed in WG2 N4925 "Naming conventions for larger character sets" we do not consider it appropriate to use algorithmic character names for

Shuishu as the repertoire is relatively small (486 characters) and there are established readings for the characters.

Proposed change by UK:

None.

Noted

An important part of the Japanese comment relates to the lack of stability for the names of the Shuishu which are classified in N4925 in the category of '1. Meaningful names'. Unlike the other 2 categories described in the same document (2. Catalog names, and 3. Algorithmic names), meaningful names inherently suffer from lack of consensus and therefore tend to be instable. Per ISO/IEC 10646 and Unicode name policy, character names are immutable once encoded. Correcting names requires the use of 'formal aliases' which is cumbersome at best. In many instances it is easier to use Algorithmic names for the formal character names and to document the 'meaningful names' in an ancillary file. This was done in the past for Nūshu.

This deserves additional discussion and therefore the encoding of Shuishu should be postponed.

GE.3. Clause 33 – Supplemental Symbols and Pictographs - 1F93F BILLIARD GAMES

We affirm our support for either of the following two options to fix the redefinition of U+1F3B1 BILLIARDS as a "Magic Eight Ball" by vendors:

A) Encode a new BILLIARD GAMES character; or

B) Revert the glyph for U+1F3B1 BILLIARDS to show a cue and a frame of balls, and add a new character called MAGIC EIGHT BALL with an "eight ball" glyph to satisfy vendor and user demand for an Eight Ball symbol as well as a Billiards symbol.

Proposed change by UK:

See above.

Noted

See also comment T8 from Ireland and TE7 from USA and its disposition.

Not every vendor has changed its glyph. According to <http://unicode.org/emoji/charts/full-emoji-list.html> Samsung still uses the symbol currently proposed at 1F93F for 1F3B1. Therefore, it may be premature to encode another character to disunify a symbol which is currently ambiguous.

GE.4. Clause 33 – Supplemental Symbols and Pictographs - 1F9A1 BADGER, 1F9A2 SWAN, 1F9A3 MAMMOTH, 1F9A4 DODO, 1F9A5 SQUIRREL

We affirm our support for including all five of these characters in Amd. 2.

Proposed change by UK:

None.

Noted

See comment TE.8 (concerning MAMMOTH and DODO) and TE.9 (SQUIRREL) from USA and their disposition.

There is no controversy concerning the BADGER and the SWAN.

ED.3. Clause 33 – Supplemental Symbols and Pictographs - 1F9B9 SUPERVILAIN

Character name is spelled incorrectly.

Proposed change by UK:

Correct spelling of character name to SUPERVILLAIN.

Accepted

See also comment T13 from Ireland and its disposition.

Strictly speaking, this should have been a technical comment (a name change is technical)

ED.4. Clause 33 – Chess Symbols - 1FA67 XIANGQI BLACK GENERAL

The tone for the alias is incorrect.

Proposed change by UK:

Correct alias from hēi jiāng to hēi jiàng.

Accepted

See also comment E1 from China and its disposition.

TE.2. Clause 33 – CJK Unified Ideographs Extension G, U+30264

U+30624 GZ-0872302: 𣎵木窓 is unifiable with U+2C114 (𣎵木窓) in Ext. E.

Proposed change by UK:

Remove U+30624 from the amendment.

Propose Accepted

See also comment T1 from China, TE.1 from Japan.

These comments will be reviewed by IRG50.

ED.5. Clause 33 – CJK Unified Ideographs Extension G, U+30C13, U+31143, U+312E1

U+30C13 (GZ-2931504): glyph is incorrect, should be 𪛗脱吝 not 𪛗脱各.

U+31143 (UK-02310): top right of glyph does not match the evidence, so the glyph should be corrected.

U+312E1 (UK-02483): the glyph does not match the evidence, and the last stroke of the left component should be corrected.

Proposed change by UK:

Correct glyphs for U+30C13, U+31143, and U+312E1.

Proposed Accepted in principle

See also comment T1 from China, TE.1 from Japan.

These comments will be reviewed by IRG50.

TE.3. Clause 33 – CJK Unified Ideographs Extension G, Additions

A number of characters submitted for IRG Working Set 2015 were excluded from the Ext. G repertoire submitted to WG2 simply because reviewers had asked questions about them. In particular 44 Zhuang-use characters were postponed even though China had provided responses to the questions asked (see IRGN2223). We believe that 39 out of these 44 pending characters do not have any outstanding issues, and meet the requirements for encoding. We therefore request that.

Proposed change by UK:

Add the following 39 characters from WS2015 to Ext. G in Amd. 2:

05276 GLGYJ-0002	02303 GZ-0802505	01513 GZ-3371101
00021 GPGLG-1003	01966 GZ-0942409	04387 GZ-3421101
03316 GPGLG-2015	01820 GZ-0952501	00064 GZ-3472102
03280 GPGLG-2019	00033 GZ-0971302	00055 GZ-3742301
05301 GPGLG-3049	00031 GZ-1252504	02638 GZ-3882101
05295 GPGLG-3050	00032 GZ-1552303	04389 GZ-4161401
05291 GPGLG-3051	02281 GZ-1571402	00017 GZ-4262303
05311 GPGLG-3052	00038 GZ-2161301	01151 GZ-4292303
03244 GPGLG-3053	01298 GZ-2481501	04597 GZ-4322205
05368 GPGLG-3056	00022 GZ-2542101	02278 GZ-4622701
03305 GPGLG-3057	01101 GZ-2562401*	00066 GZ-4632101
03776 GZ-0591501	04616 GZ-2971101	04872 GZ-4871501
03561 GZ-0601304	02454 GZ-3141602	00030 GZ-4921301

*For GZ-2562401, we recommend normalizing the glyph form to 巷网 as this the expected glyph according to Chinese conventions.

Propose Accepted in Principle

See also comment T1 from China, TE.1 from Japan.

These comments will be reviewed by IRG50.

USA: Negative

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

TE.1. Latin Extended-D U+A7C0, U+A7C1

The USNB requests the removal of U+A7C0 LATIN CAPITAL LETTER THORN WITH DIAGONAL STROKE and U+A7C1 LATIN SMALL LETTER THORN WITH DIAGONAL STROKE. Both these are glyph variants, and can be represented with OpenType features or different fonts.

Proposed change by USA:

Remove U+A7C0 LATIN CAPITAL LETTER THORN WITH DIAGONAL STROKE and U+A7C1 LATIN SMALL LETTER THORN WITH DIAGONAL STROKE.

Proposed Accepted

See also comments T2 from Ireland

This needs to be further reviewed and therefore will be postponed.

TE.2. Nandinagari (addition of U+11BD4)

The USNB requests U+11BD4 NANDINAGARI VOWEL SIGN PRISHTHAMATRA E be added to the PDAM, with glyph and properties as shown in the proposal WG2 N4928 (UTC [L2/17-213r](#)), since the Nandinagari script is already under PDAM ballot.

Proposed change by USA:

Add U+11BD4 NANDINAGARI VOWEL SIGN PRISHTHAMATRA E.

Proposed Accepted

If the comment TE.3 from USA below is accepted, the code point becomes 119E4.

TE.3. Nandinagari (block shift)

The USNB requests Nandinagari be shifted from the range U+11B90..U+11BEF to the range U+119A0..U+119FF, as currently shown in the SMP roadmap.

Proposed change by USA:

Move Nandinagari to U+119A0..U+119FF.

Proposed Accepted

This impact comment TE.2 above.

TE.4. Shuishu Logograms and Shuishu Radicals

The USNB requests Shuishu Logograms and Shuishu Radicals be removed from the PDAM ballot. This large pictographic set should have an established record of publication and stronger indication that the repertoire is complete and well-analyzed before it is standardized.

Proposed change by USA:

Remove Shuishu Logograms and Shuishu Radicals. If this comment is satisfied, the vote will be changed to "Yes".

Proposed Accepted in principle

See also comments T6 from Ireland, and comment TE.2 from Japan and its disposition.

An important part of the Japanese comment relates to the lack of stability for the names of the Shuishu which are classified in N4925 in the category of '1. Meaningful names'. Unlike the other 2 categories described in the same document (2. Catalog names, and 3. Algorithmic names), meaningful names inherently suffer from lack of consensus and therefore tend to be instable. Per ISO/IEC 10646 and Unicode name policy, character names are immutable once encoded. 'Correcting names' requires the use of 'formal aliases' which is cumbersome at best. In many instances it is easier to use Algorithmic names for the formal character names and document the 'meaningful names' in an ancillary file. This was done in the past for Nūshu.

This deserves additional discussion and therefore the encoding of Shuishu should be postponed.

TE.5. Supplemental Punctuation (U+2E4F)

The USNB requests U+2E4F DOUBLE OBLIQUE HYPHEN WITH FALLING DOTS be renamed MEDIEVAL CORNISH VERSE SEPARATOR, based on the typography. More justification needs to be provided to determine the representative glyph for this character.

Proposed change by USA:

Change the name of U+2E4F to MEDIEVAL CORNISH VERSE SEPARATOR.

Proposed Accepted

See also comment E2 from Ireland and its disposition.

TE.6. Khitan Small Script

The USNB requests the removal of Khitan Small Script until the model can be settled. No joiners are included in the PDAM, so the model is not complete.

Proposed change by USA:

Remove the Khitan Small Script.

Proposed Accepted in principle

See also comment T5 from Ireland and its disposition.

The encoding of Khitan Small Script should be postponed.

TE.7. Supplemental Symbols and Pictographs - 1F93F BILLIARD GAMES

The USNB requests the removal of U+1F93F BILLIARD GAMES, and requests the glyph for U+1F3B1 BILLIARDS be changed back to the original glyph (below).



1F3B1

With the change of glyph for U+1F3B1 BILLIARDS back to the original shape (with cue stick and balls), the character can be used for “billiard games” so no new character is needed.

Proposed change by USA:

Remove U+1F93F BILLIARD GAMES and change the glyph of U+1F3B1 BILLIARDS.

Proposed Accepted

See also comment T8 from Ireland and GE3 from UK and their disposition.

Not every vendor has changed its glyph. According to <http://unicode.org/emoji/charts/full-emoji-list.html>

Samsung still uses the symbol currently proposed at 1F93F for 1F3B1. Therefore, it may be premature to encode another character to disunify a symbol which is currently ambiguous.

TE.8. Supplemental Symbols and Pictographs - 1F9A3 MAMMOTH, 1F9A4 DODO

The USNB requests U+1F9A3 MAMMOTH and U+1F9A4 DODO be removed from the PDAM. These characters do not meet the selection requirements for adding new emoji nor for adding pictographic symbols to the standard.

Proposed change by USA:

Remove MAMMOTH and DODO.

Proposed Accepted in principle

See also comment T11 from Ireland and GE4 from UK and their disposition.

The encoding of MAMMOTH and DODO is postponed.

TE.9. Supplemental Symbols and Pictographs - 1F9A5 SQUIRREL

The USNB requests the removal of U+1F9A5 SQUIRREL, and an annotation “squirrel” be added to U+1F43F CHIPMUNK. Squirrel does not look sufficiently distinctive in small, emoji sizes. This character does not meet the selection requirements for adding new emoji nor for adding pictographic symbols to the standard.

Proposed change by USA:

Remove SQUIRREL and add an annotation “squirrel” to U+1F43F CHIPMUNK.

Proposed Accepted

See also comment T11 from Ireland and GE4 from UK and their disposition.

Looking at <http://unicode.org/emoji/charts/full-emoji-list.html> it is clear that there is ambiguity at the representation of 1F43F CHIPMUNK and many of the chosen glyphs already look like Squirrel. Chipmunk and Squirrel are two species that are closely related (even more so the ground Squirrel and the Squirrel).

The proposed new annotation for 1F43F would be:

- Also used to represent a squirrel

TE.10. Supplemental Symbols and Pictographs - 1F97B TROLL

The USNB requests the removal of U+1F97B TROLL. No proposal for this character was submitted. At emoji sizes, it is not differentiated clearly from U+1F479 JAPANESE OGRE and U+1F47A JAPANESE GOBLIN. This character does not meet the selection requirements for adding new emoji nor for adding pictographic symbols to the standard.

Proposed change by USA:

Remove TROLL.

Proposed Accepted

See also comment T10 from Ireland and its disposition.

The encoding of TROLL is postponed.

TE.11. Supplemental Symbols and Pictographs - 1F9F1 BRICKS

The USNB requests the character name for U+1F9F1 BRICKS be changed to BRICK with a new glyph. BRICK has clear metaphorical usage which BRICKS does not.

Proposed change by USA:

Change the name and glyph for U+1F9F1 to BRICK.

Proposed Accepted

See also comment E7 from Ireland and its disposition.

TE.12. Supplemental Symbols and Pictographs - 1F9FB ROLL OF TOILET PAPER

The USNB requests the name for U+1F9FB ROLL OF TOILET PAPER be changed to ROLL OF PAPER. The new name is more general, and hence preferable.

Proposed change by USA:

Change the name of the character of U+1F9FB to ROLL OF PAPER.

Proposed Accepted

TE.13. Wancho – Tone characters

The USNB requests the removal of the following four tone characters in Wancho:

1E2EC WANCHO TONE TUP

1E2ED WANCHO TONE TUP MANG

1E2EE WANCHO TONE OKOI

1E2EF WANCHO TONE OKOI MANG

Evidence of the characters in print is needed to justify these characters.

Proposed change by USA:

Remove 4 Wancho tone characters.

Proposed Accepted

See also comment T7 from Ireland and its disposition.

The encoding of these 4 characters is postponed, this may be added back in Amendment 2 assuming further evidence.

TE.14. Egyptian Hieroglyph Format Controls – Glyph updates

The USNB request that the following 4 characters have updated glyphs:

U+13432 EGYPTIAN HIEROGLYPH START AT TOP

U+13433 EGYPTIAN HIEROGLYPH START AT BOTTOM

U+13434 EGYPTIAN HIEROGLYPH END AT TOP

U+13435 EGYPTIAN HIEROGLYPH END AT BOTTOM.

For the glyphs, we recommend using small tile-shaped solid symbols representing the corner positions inside a dotted box, but without abbreviation letters.

For the glyphs of U+13430..U+13431 and U+13436..U+13438, we support retaining the glyphs as shown in the amendment.

Proposed change by USA:

Update the glyphs U+13432..U+13435 as described, and retain the glyphs for U+13430..U+13431 and U+13436..U+13438, as shown in the amendment.

Proposed Accepted

See also comment E5 from Ireland and its disposition.

TE.15. CJK Extension G

The USNB has noted some radical stroke anomalies in Extension G, and as a result we request Extension G not be progressed to enquiry ballot until the USNB has had an opportunity to complete its review of this large repertoire. The IRG is still reviewing Extension G.

Proposed change by USA:

See above.

Proposed Accepted

See also comment T1 from China, TE.1 from Japan. TE.2, TE.3 , and ED.5 from UK and their disposition.

All these comments should be considered by IRG 50 and result in a new version of CJK Extension which may or may not be brought back into a future DAM2 based on discussion by experts and SC2 members at the next SC2 and WG2 meeting in London (June 2018).