

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

Doc. Type: Draft disposition of comments

Title: Draft disposition of comments on SC2 N 4049 (FPDAM text for Amendment 6 to ISO/IEC 10646:2003)

Source: Michel Suignard (project editor)

Project: JTC1 02.10646.00.06

Status: For review by WG2

Date: 2009-04-15

Distribution: WG2

Reference: SC2 N4049, N4065, WG2 N3599

Medium: Paper, PDF file

Comments were received from the following P-Members: Germany, Ireland, Japan, UK, and USA, in SC2 N4049 (Summary of Voting/Tables of Replies) and from another P-Member: India in WG2 N3599. The following document is the disposition of those comments. The disposition is organized per country in reverse alphabetical order (done this time because many US comments were cross referenced by other NBs)

Note – The full content of the ballot comments (minus some figures and charts) 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.

USA: Positive with comments:

The U.S. National Body is voting Yes with comments on the following SC2 ballot: SC2 N4049: Information technology -- Universal Multiple-Octet Coded Character Set (UCS) --AMENDMENT 6: Javanese, Lisu, Meetei Mayek, Samaritan, and other characters.

Technical Comments:

T.1 Vedic Extensions – renaming for 1DC4

The U.S. requests the following name change be made in the Vedic Extensions block: Change U+1CD4 from VEDIC SIGN YAJURVEDIC KASHMIRI SVARITA to VEDIC SIGN MIDCHARACTER SVARITA. This change is requested based on input from the Government of India, whose experts said the character may appear in manuscripts outside Kashmir. The U.S. expert who co-authored the proposal is agreeable to this change.

WG2 discussion

See similar comment from Germany T.1 and Ireland T.4.

*US and Germany support the following change for 1DC4 VEDIC SIGN KASHMIRI SVARITA:
VEDIC SIGN MIDCHARACTER SVARITA*

Ireland proposes:

VEDIC SIGN YAJURVEDIC MIDLINE SVARITA

The main point is to determine whether or not Yajurvedic should appear in the name (MIDLINE should be acceptable to all parties). This could also be captured in annotation one way or the other.

T.2 Enclosed Alphanumeric Supplement – renaming of some characters

a) The U.S. requests the following ten names of symbols contained in Enclosed Alphanumeric Supplement block be changed, with the sub-headers modified as appropriate. The original names of these character contained “WHITE ON BLACK”, a name not found elsewhere in the UCS. The term “NEGATIVE” is customary usage for such white-on-black characters (cf. U+24EB NEGATIVE CIRCLED NUMBER ELEVEN), and we recommend it be adopted.

- U+1F157 WHITE ON BLACK CIRCLED LATIN CAPITAL LETTER H should be changed to: NEGATIVE CIRCLED LATIN CAPITAL LETTER H
- U+1F15F WHITE ON BLACK CIRCLED LATIN CAPITAL LETTER P should be changed to: NEGATIVE CIRCLED LATIN CAPITAL LETTER P
- U+1F179 WHITE ON BLACK SQUARED LATIN CAPITAL LETTER J should be changed to: NEGATIVE SQUARED LATIN CAPITAL LETTER J
- U+1F17B WHITE ON BLACK SQUARED LATIN CAPITAL LETTER L should be changed to: NEGATIVE SQUARED LATIN CAPITAL LETTER L
- U+1F17C WHITE ON BLACK SQUARED LATIN CAPITAL LETTER M should be changed to: NEGATIVE SQUARED LATIN CAPITAL LETTER M
- U+1F17F WHITE ON BLACK SQUARED LATIN CAPITAL LETTER P should be changed to: NEGATIVE SQUARED LATIN CAPITAL LETTER P
- U+1F18A CROSSED WHITE ON BLACK SQUARED LATIN CAPITAL LETTER P should be changed to: CROSSED NEGATIVE SQUARED LATIN CAPITAL LETTER P
- U+1F18B WHITE ON BLACK SQUARED IC should be changed to: NEGATIVE SQUARED IC
- U+1F18C WHITE ON BLACK SQUARED PA should be changed to: NEGATIVE SQUARED PA
- U+1F18D WHITE ON BLACK SQUARED SA should be changed to: NEGATIVE SQUARED SA

Propose acceptance

Also supported by comment T.2a) from Germany and T.9.a from Ireland.

b) The U.S. also requests character U+1F14F SQUARE DJ be moved to U+1F190. This change is requested because characters in the U+1F14x column are Latin letters enclosed in squares. Since this character is not contained in a square, it should be moved out of the U+1F14x column.

Propose acceptance

Also supported by comment T.2b) from Germany.

T.3 Enclosed Alphanumeric Supplement – addition of WZ

The U.S. requests the following character (with revised name) be added to this amendment:

U+1F12E CIRCLED WZ. The rationale, glyph, and decomposition are provided in document L2/09-034 [N3579].

Propose acceptance

Also supported by comment T.3 from Germany.

T4. Unified Canadian Aboriginal Syllabics Extended block - reordering

The U.S. recommends the characters in this block be reordered as requested in document N3563[read N3533].

Propose acceptance

Also supported by comment T.4 from Germany with the corrected document number. Also supported by comment T.3 from Ireland.

T5. Kaithi – addition of 2 characters

The U.S. requests two characters be added to the Kaithi block: U+110BE = KAITHI SECTION MARK and U+110BF = KAITHI DOUBLE SECTION MARK, as documented in L2/09-089 (N3574).

Propose acceptance

Also supported by comment T.5 from Germany and T.8 from Ireland.

Editorial Comments:

E1. Devanagari Extended

The U.S. requests the annotation “used for insertion of characters” be added to the character U+A8FA DEVANAGARI CARET, based on feedback received from the Government of India.

Accepted

E2. Old South Arabian

The U.S. requests the addition of annotations for the three “s” sounds in Old South Arabian to clarify that the current character names do not reflect the etymological relations for the sounds in old and modern-day Hebrew.

For U+10A66 OLD SOUTH ARABIAN LETTER SHIN, add the annotation: “the sound represented by this letter is cognate with Classical and Modern Hebrew sin”

For U+10A6A OLD SOUTH ARABIAN LETTER SAMEKH, add the annotation: “the sound represented by this letter is cognate with Classical and Modern Hebrew shin”

For U+10A6F OLD SOUTH ARABIAN LETTER SIN, add the annotation: “the sound represented by this letter is cognate with Classical and Modern Hebrew samekh”

WG2 discussion

Although Editorial comments are non technical by nature, these are cross referenced by technical comment T.7 from Ireland which requires a WG2 discussion, including name changes.

E3. Miscellaneous Symbols

The U.S. requests the glyph be improved for U+26DF BLACK TRUCK so that its size is comparable to similar symbols in the block.

Accepted

UK, Positive with comments

Editorial comments

E.1 Clause 34, Row 0F0 Tibetan, cross-reference for 0FCC

0FCC cross-references to 534D, but should cross-reference to 0FD6 as indicated on page 4. Change cross-reference for 0FCC to 0FD6.

Accepted

Due to the chart production process, it is sometimes difficult to represent modified annotation to existing characters until the non technical ballot (FDAM). This is the reason why they are also captured in the main body of the amendment.

E.2 Clause 34, Row 0F0 Tibetan, 0FD5-0FD8 header

0FD5..0FD8 are headed "Symbols used in Buddhism and Hinduism", but some are also sacred symbols for the Bön and Jainist religions. Change heading to "Symbols used in Bön, Buddhism, Hinduism and Jainism" or simply "Religious Symbols".

Accepted

It will use 'Religious Symbols'.

E.3 Clause 34, Row 1F20 Enclosed Ideographic Supplement

The annotations for 1F243 and 1F248 do not make sense, and do not appear to reflect the meaning of these characters. Change annotation for 1F243 to "safe", and annotation for 1F248 to "steal".

Accepted

See also comment E8 from Ireland.

Japan, Negative

Technical comments

JPT1- Enclosed CJK Letters and Months block (3200-32FF) – Enclosed katakatana DE

The symbol ARIB-9056 that has a shape of "katakana de enclosed in square" is allocated at U+32FF. Japan believes the allocation is not best. The concerns include: (1) the position looks best for a future addition of a new katakana letter enclosed in a circle, e.g., CIRCLED KATAKANA N, or some other letter enclosed in a circle or parentheses; and (2) there is no square-enclosed symbol in this block.

Allocate the character currently allocated at U+32FF in somewhere else. The suggested new block is Enclosed Ideographic Supplement (1F200-1F2FF).

WG2 discussion

Strictly speaking, 'enclosed' can mean parenthesized, circled, or squared, as used in many other blocks. Furthermore the enclosed character is a katakana character like the previous characters in the block. The editor has no strong preferences one way or another but does not see the current location as problematic. This is also related to the resolution of comment JPT3 as accepting it would probably means that this character would be inserted after 'two-way broadcasting service' (like in the original set).

JPT2: Codechart for Enclosed Ideographic Supplement (1F200-1F2FF) – Squared 4E8C

In the ARIB STD-B24, there are two similar symbols ARIB-9058 and ARIB-9330 that have shapes "an ideograph two enclosed in a square". In FPDAM 6, however, there is only one graphic symbol of such shape: 1F213. Japan worries about possibility of roundtrip integrity issues and wants WG2 to discuss this issue in WG2 meeting before approves the FPDAM.

Allocate two similar symbols at two separate code positions.

WG2 discussion

Note that the single proposed character has already a compatibility decomposition to 4E8C. A new character would have the same decomposition. The only difference in the original ARIB set is that there are two semantic meanings allocated to inherently the same character. This is a common occurrence in writing systems and is typically not preserved in 10646 encoding. Strict source separation rule was not part of encoding the Arib set into 10646. For example, size variation was not used to differentiate sources.

JPT3: Enclosed Ideographic Supplement (1F200-1F2FF) – Ordering

Currently, the block arranges characters in a order of UCS code positions of enclosed ideographs, after grouping by the enclosing shapes. It can confuse users. In the ARIB STD-B24, the characters are grouped and ordered in the intended usages or purposes, and the ordering seems better.

Rearrange the characters in Enclosed Ideographic Supplement block following the order in ARIB STD-B24.

Propose acceptance

Because new additions to this block would not follow the order of the original UCS code positions, it seems reasonable to accept this request. The block sub-headers in the charts already capture the ARIB sources. The new order would be as follows (previous code points as in FPDAM6 still used as identifiers, the first group would start at 1F210, the second at 1F240, although it could be argued that they should be no hole between the two groups):

```
@           Squared ideographs from ARIB STD B24
1F221 SQUARED CJK UNIFIED IDEOGRAPH-624B
      = broadcasting service with sign language interpretation
      # <square> 624B
1F21E SQUARED CJK UNIFIED IDEOGRAPH-5B57
      = close-captioned broadcasting
      # <square> 5B57
1F218 SQUARED CJK UNIFIED IDEOGRAPH-53CC
      = two-way broadcasting service
      # <square> 53CC
32FF  SQUARED KATAKANA DE
      = data broadcasting service linked with a main program
```

<square> 30C7
 1F213 SQUARED CJK UNIFIED IDEOGRAPH-4E8C
 = bilingual broadcasting service
 = second base
 # <square> 4E8C
 1F21C SQUARED CJK UNIFIED IDEOGRAPH-591A
 = sound-multiplex broadcasting service
 # <square> 591A
 1F22D SQUARED CJK UNIFIED IDEOGRAPH-89E3
 = commentary broadcasting
 # <square> 89E3
 1F21D SQUARED CJK UNIFIED IDEOGRAPH-5929
 = weather forecast
 # <square> 5929
 1F214 SQUARED CJK UNIFIED IDEOGRAPH-4EA4
 = traffic information
 # <square> 4EA4
 1F228 SQUARED CJK UNIFIED IDEOGRAPH-6620
 = drama film
 # <square> 6620
 1F22A SQUARED CJK UNIFIED IDEOGRAPH-7121
 = free broadcasting service
 # <square> 7121
 1F226 SQUARED CJK UNIFIED IDEOGRAPH-6599
 = pay broadcasting
 # <square> 6599
 1F217 SQUARED CJK UNIFIED IDEOGRAPH-524D
 = the first part
 # <square> 524D
 1F220 SQUARED CJK UNIFIED IDEOGRAPH-5F8C
 = the latter part
 # <square> 5F8C
 1F215 SQUARED CJK UNIFIED IDEOGRAPH-518D
 = rebroadcast
 # <square> 518D
 1F227 SQUARED CJK UNIFIED IDEOGRAPH-65B0
 = new series of programme
 # <square> 65B0
 1F216 SQUARED CJK UNIFIED IDEOGRAPH-521D
 = first released program
 # <square> 521D
 1F22C SQUARED CJK UNIFIED IDEOGRAPH-7D42
 = the last episode
 # <square> 7D42
 1F22B SQUARED CJK UNIFIED IDEOGRAPH-751F
 = live broadcast
 # <square> 751F
 1F22E SQUARED CJK UNIFIED IDEOGRAPH-8CA9
 = mail-order
 # <square> 8CA9
 1F21B SQUARED CJK UNIFIED IDEOGRAPH-58F0
 = voice actors
 # <square> 58F0
 1F21A SQUARED CJK UNIFIED IDEOGRAPH-5439
 = dubbed version
 # <square> 5439
 1F229 SQUARED CJK UNIFIED IDEOGRAPH-6F14
 = performed by
 # <square> 6F14
 1F223 SQUARED CJK UNIFIED IDEOGRAPH-6295
 = pitcher
 # <square> 6295
 1F225 SQUARED CJK UNIFIED IDEOGRAPH-6355
 = catcher
 # <square> 6355
 1F210 SQUARED CJK UNIFIED IDEOGRAPH-4E00
 = first base
 # <square> 4E00
 1F211 SQUARED CJK UNIFIED IDEOGRAPH-4E09
 = third base
 # <square> 4E09
 1F230 SQUARED CJK UNIFIED IDEOGRAPH-904A
 = short stop
 # <square> 904A

```

1F21F SQUARED CJK UNIFIED IDEOGRAPH-5DE6
= left field
# <square> 5DE6
1F212 SQUARED CJK UNIFIED IDEOGRAPH-4E2D
= center field
# <square> 4E2D
1F219 SQUARED CJK UNIFIED IDEOGRAPH-53F3
= right field
# <square> 53F3
1F224 SQUARED CJK UNIFIED IDEOGRAPH-6307
= designated hitter
# <square> 6307
1F22F SQUARED CJK UNIFIED IDEOGRAPH-8D70
= runner
# <square> 8D70
1F222 SQUARED CJK UNIFIED IDEOGRAPH-6253
= batter
# <square> 6253
@ Ideographs with tortoise shell brackets from ARIB STD B24
1F246 TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-672C
= home run
# <compat> 3014 672C 3015
1F240 TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-4E09
= triple
# <compat> 3014 4E09 3015
1F241 TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-4E8C
= double
# <compat> 3014 4E8C 3015
1F243 TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-5B89
= safe
# <compat> 3014 5B89 3015
1F247 TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-70B9
= run
# <compat> 3014 70B9 3015
1F244 TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-6253
= hit
# <compat> 3014 6253 3015
1F248 TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-76D7
= steal
# <compat> 3014 76D7 3015
1F242 TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-52DD
= win
# <compat> 3014 52DD 3015
1F245 TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-6557
= loss
# <compat> 3014 6557 3015

```

JPT4: Enclosed Ideographic Supplement (1F200-1F2FF) – Naming

The block for U+1F200 to U+1F2FF is currently called "Enclosed Ideographic Supplement." Although existing block of a name "XXX supplement" has a corresponding block of name "XXX", there is no such block of the name "Enclosed Ideographic." Considering the nature of the allocated symbols, this block seems a supplement to the existing "Enclosed CJK Letters and Months" block.

Change the name of the block for 1F200-1F2FF to "Enclosed CJK Letters and Months Supplement".

WG2 discussion

There are no strict naming rules for 'extension' to existing blocks. Supplement and Extensions are used, and because block names are stable, various naming schemes have been used through various editions and amendments of the standard and not always consistently. Some examples of inconsistencies:

- *Cyrillic, Cyrillic Supplement, Cyrillic Extended-A, Cyrillic Extended-B*
- *Basic Arabic, Arabic Supplement*
- *KangXi Radicals, CJK Radicals Supplement*
- *Enclosed Alphanumerics, Enclosed Alphanumeric Supplement,*
- *Enclosed CJK Letters and Months, Enclosed Ideographic Supplement*

Using 'Enclosed CJK Letters and Months Supplement' seems a bit misleading because there are no CJK Months in the new block and will likely never be. Maybe another term such as 'Enclosed CJK Letters Supplement' could be used.

JPT5: Clause 3 Normative references – Unicode 5.2

The draft contains updates to normative references to two unpublished Unicode documents dated Oct. 2009. Japan can't judge whether it is appropriate to make the new version normative parts of ISO/IEC 10646 before knowing the changes from the current versions.

Japan wants Unicode Consortium to submit the draft for 5.2.0 of the referred documents (or summary of planned changes, at least) to WG2 for evaluation by members.

WG2 discussion

This is a concern that has been mentioned several times by Japan and can be addressed in two different ways:

- 1. Use an existing version of the Unicode Standard referenced document. The problem with that is that it creates a lack of synchronization between the two standards, because the reference would point to an earlier version of Unicode than the one covering the same repertoire.*
- 2. Provide to WG2 drafts of the referred documents as suggested by Japan.*

The editor is in favor of the second solution. Note that at this moment, draft updates are available for both UAX#9 and UAX#15 at <http://www.unicode.org/reports/tr9/tr9-20.html> <http://www.unicode.org/reports/tr15/tr15-30.html> respectively. It is a matter for the Unicode Consortium to establish a policy to come with a draft version of the normative references during the technical balloting of the amendments including these references.

Editorial comments

JPE6: Clause 29 Named UCS Sequence Identifiers - Tamil

The order of the new TAMIL USIs in the list seems inconsistent. In particular, the order the USI <0B95, 0BCD, 0BB7, 0BCD> TAMIL CONSONANT KSS seems strange. Unless there is a good reason to list them in the current ordering, it should be listed in a more consistent manner.

List them in a so-called dictionary ordering based on the code positions of USIs. (Any other consistent ordering is fine.)

Propose acceptance in principle

There is in fact a consistency through the list. The first section <0B95, 0BCD>..<0B95, 0BCD, 0BB7, 0BCD> is the list of all 24 consonants without the inherent vowel (using the virama 0BCD to suppress the inherent vowel). Following are all the same consonants in the same order but associated with all non inherent vowels (11). The last group is a bit special as it also includes the inherent vowel combination KSSA because it is made internally of a consonant cluster (KSS), for a total of 12 sequences. The very last sequence is the special case SHRII. In other words, the order of the sequences should make sense to users of these Tamil letters.

JPE7: Codechart for Enclosed Alphanumeric Supplement – 1F12D

At the code position 1F12D, the current code chart shows a flat ellipsis around "CD". It looks strange when compared to the chart in ARIB STD-B24 and may confuse users.

Use a round circle around "CD".

Accepted in principle

Pending access to an improved glyph. See also comment T.7 from Germany.

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 30-1348, Code Tables and list of character names and New blocks.

If T3 below is accepted, add “18C6-18DF, 18F3-18F5” in FDAM 6.

Propose acceptance in principle

Note that the original range 18E0-18FC was missing from the code position enumeration on page 5 of FPDAM6. However, based on other comments by USA and Germany (T.4 for both NBs), it is likely that the whole repertoire proposed between Amendment 6 and 7 will be added in one step in Amendment 6. In that case, the added range will be 18B0-18F5.

T2. Page 6, Annex A.2.1. UCAS Extended

For “United Canadian Aboriginal Syllabics Extended-A”, read “Unified Canadian Aboriginal Syllabics Extended”.

Accepted

T.3 Page 38, Row 18B: Unified Canadian Aboriginal Syllabics Extended.

With reference to ISO/IEC JTC1/SC2/WG2 N3533R “Proposed Revision for Unified Canadian Aboriginal Syllabics Extended Block”, Ireland requests that the characters shown in yellow in FPDAM 6 be rearranged as shown in the second of the three charts in that document. Ireland is confident in the technical content of this repertoire and would not oppose the transfer of the Canadian Syllabics characters in PDAM 7 to FDAM 6, but does not insist on such a transfer.

Propose acceptance

Also supported by comments T.4 from both Germany and USA.

T.4 Page 42, Row 1CD: Vedic Extensions.

Ireland requests that the name of 1CD4 be changed from VEDIC SIGN YAJURVEDIC KASHMIRI SVARITA to VEDIC SIGN YAJUR VEDIC MIDLINE SVARITA.

WG2 discussion

See disposition of comment T.1 from USA. (Related comment T.1 from Germany)

T.5 Page 81, Row AA6: Myanmar Extended-A.

With reference to ISO/IEC JTC1/SC2/WG2 N3594 “Proposal for encoding one additional Myanmar character for Pa’o Karen in the UCS”, Ireland requests that a new character AA7B MYANMAR SIGN PAO KAREN TONE be added to FDAM 6.

WG2 discussion

T.6 Page 82, Row ABC: Meetei Mayek.

With reference to ISO/IEC JTC1/SC2/WG2 N3478R “Proposed Encoding for Meetei Mayek Extended Block”, Ireland requests that a new block “Meetei Mayek Extensions” beginning at AAE0 be added to FDAM 6.

Propose out of scope

The amendment contains no character part of that block, so there is no point in creating an empty block in 10646 before some characters are accepted by WG2 for balloting.

T.7. Page 87, Row 10A6: Old South Arabian.

Ireland would like to clarify some things about three sibilants under ballot, in anticipation of editorial comments we have seen to be proposed by the US National Body. There are two issues here. The first is the genetic relationship of the graphic characters, and the second is the phonetic use to which they are put.

10A66 𐤱 OLD SOUTH ARABIAN LETTER SHIN, transliterated š or s₂, representing [ʃ], corresponds in a genetic relationship directly to:

- 10914 𐤱 PHOENICIAN LETTER SHIN, [ʃ]
- 1220 𐌪 ETHIOPIC SYLLABLE SZA or śawt, Ge‘ez [ʃ], Amharic [s]
- 05E9 𐤱 HEBREW LETTER SHIN, represents three sounds (according to Richard Goerwitz) in reconstructed mid-2nd Millennium Hebrew *[ʃ ʃ s], and two sounds in Tiberian Hebrew [ʃ s] and in Modern Hebrew [ʃ s].

Ireland proposed to name this character according to its genetic relationship to characters in other Semitic scripts. Since the reconstructed original sound represented by this graphic characters is *[ʃ], we are not quite sure why the US National Body would wish to annotate this character as “the sound represented by this letter is cognate with Classical and Modern Hebrew sin”, since what we have is a three-way proto-Semitic distinction *[ʃ ʃ s] falling together as [ʃ s] in Hebrew, with (pointed) 𐤱 representing [ʃ] and (pointed) 𐤱 and 𐤎 representing [s]. In any case the annotation seems imprecise. One might say “the sound represented by this letter is cognate with the sound represented by Classical and Modern Hebrew sin”, but then the sound is likewise cognate with sounds in other languages, and Hebrew samekh and sin both represent the same sound, [s].

10A6A 𐤱 OLD SOUTH ARABIAN LETTER SAMEKH, transliterated s or s₁, representing [s], corresponds in a genetic relationship directly to:

- 1090E 𐤱 PHOENICIAN LETTER SEMK, [s]
- 1230 𐌪 ETHIOPIC SYLLABLE SZA or sāt, Ge‘ez [s], Amharic [s]
- 05E9 𐤱 HEBREW LETTER SAMEKH, reconstructed mid-2nd Millennium Hebrew [ts], Tiberian Hebrew [s], Modern Hebrew [s]

Ireland proposed to name this character according to its genetic relationship to characters in other Semitic scripts. Since the reconstructed original sound is *[s], we are not quite sure why the US National Body would wish to annotate this character as “the sound represented by this letter is cognate with Classical and Modern Hebrew shin”, since Phoenician, Ge‘ez, Amharic, and Hebrew all have [s] for samekh.

10A6F 𐤱 OLD SOUTH ARABIAN LETTER SIN, transliterated ś or s₃, representing dental [s̪], has no genetic character correspondence with Phoenician, Ethiopic, or Hebrew. (The character looks, superficially at least, like two back-to-back 𐤱 characters.) We are not sure why the US National Body would wish to annotate this character as “the sound represented by this letter is cognate with Classical and Modern Hebrew samekh”.

WG2 discussion

Some apparent typos fixed by the editor:

2nd 10A66 changed to 10A6A, in same section 1220 𐌪 ETHIOPIC SYLLABLE SZA changed to 1230 𐌪 ETHIOPIC SYLLABLE SA. Third 10A66 changed to 10A6F.

Another mistake not fixable: 05E9 𐤱 HEBREW LETTER SAMEKH, because either the glyph or the name is wrong.

As a result of these typos, the editor is totally lost in this explanation. Related to US comment E.2 (editorial).

T7a. It seems that the US National Body comments attempt to encapsulate encyclopedic information that cannot really be encapsulated. At a minimum an explanation (with examples) is needed to determine just what the intent here is.

WG2 discussion

See above.

T7b. Ireland believes that it is both anachronistic and misleading to use the name SIN for this Old South Arabian character (since it has no real genetically correspondent character in other Semitic scripts), and, accordingly, requests that the name be changed to the neutral OLD SOUTH ARABIAN LETTER SE, as documented previously in our comments on PDAM 6.2 (2008-09-23).

WG2 discussion

See above.

T8. Page 95, Row 1108: Kaithi.

With reference to ISO/IEC JTC1/SC2/WG2 N3574 “Proposal to Encode Section Marks for Kaithi in ISO/IEC 10646”, Ireland requests that two characters be added to this block:

110BE KAITHI SECTION MARK
110BF KAITHI DOUBLE SECTION MARK

Propose acceptance

Also supported by comments T.5 from both Germany and USA.

T9. Page 97, Row 1F10: Enclosed Alphanumeric Supplement.

a) Ireland requests that the set of characters between 1F130..1F148, 1F150..1F168, and 1F170..1F188 be encoded completely and comprehensively. It is clear from the configuration in the code table what characters are intended to fill in the empty spaces, and it is clear that no other characters are expected to be encoded there. Already we have seen a request for three additions at 1F170, 1F171, and 1F17E. It is certain that the remainder of the characters will one day turn up, and it is a false economy and a waste of valuable resources to delay encoding this set as a complete set now.

WG2 discussion

Ireland is correct in expected content at these locations. It is however unclear whether we should encode characters before actual use is demonstrated for the other characters along with formal proposals. The list below also contains three characters not covered by the rationale above (WC twice and AB).

b) Ireland also requests that the letters on the ballot named WHITE ON BLACK CIRCLED LATIN LETTER and WHITE ON BLACK SQUARED LATIN LETTER be renamed to NEGATIVE CIRCLED LATIN LETTER and NEGATIVE SQUARED LATIN LETTER respectively.

Propose acceptance

Also supported by comments T.2 from both Germany and USA.

c) We also request for the following informative annotations to a number of characters as outlined here.

- 1F12E CIRCLED WZ
- 1F130 SQUARED LATIN CAPITAL LETTER A
= Anreisetag (day of arrival)
- 1F131 SQUARED LATIN CAPITAL LETTER B
= bank
- 1F130 SQUARED LATIN CAPITAL LETTER T
= tobacconist
- 1F14F SQUARED WC
= public toilet
- 1F150 NEGATIVE CIRCLED LATIN CAPITAL LETTER A
= airport
- 1F155 NEGATIVE CIRCLED LATIN CAPITAL LETTER F
= ferry
- 1F15A NEGATIVE CIRCLED LATIN CAPITAL LETTER K
= Kurort (spa)
- 1F15B NEGATIVE CIRCLED LATIN CAPITAL LETTER L
= light rail
- 1F15C NEGATIVE CIRCLED LATIN CAPITAL LETTER M
= metro
- 1F161 NEGATIVE CIRCLED LATIN CAPITAL LETTER R
= Raststätte (rest stop)
- 1F162 NEGATIVE CIRCLED LATIN CAPITAL LETTER S
= Stadtbahn (metropolitan railway)
- 1F163 NEGATIVE CIRCLED LATIN CAPITAL LETTER T
= Tankstelle (petrol station)
- 1F164 NEGATIVE CIRCLED LATIN CAPITAL LETTER U
= underground rail
- 1F170 NEGATIVE SQUARED LATIN CAPITAL LETTER A
= blood type A
- 1F171 NEGATIVE SQUARED LATIN CAPITAL LETTER B
= blood type B
- 1F172 NEGATIVE SQUARED LATIN CAPITAL LETTER C
= centre, city centre
- 1F175 NEGATIVE SQUARED LATIN CAPITAL LETTER F
= Feuerwache (fire department)
- 1F177 NEGATIVE SQUARED LATIN CAPITAL LETTER H
= hospital
- = Haltstelle (halting place)
- 1F17E NEGATIVE SQUARED LATIN CAPITAL LETTER O

- = blood type O
- 1F18E NEGATIVE SQUARED AB
= blood type AB
- 1F17F NEGATIVE SQUARED WC
= public toilet

The code table corresponding to these comments is attached below. [Please refer to SC2 N4065 for details]

WG2 discussion

See above. On this list, only 1F131 SQUARE LATIN CAPITAL LETTER B is currently proposed in amendment 6 and has already annotations that should be preserved. In addition, 1F12E CIRCLED WZ IS BEING PROPOSED by the USA and Germany (comments T.3 for both). All others are new. Note as well that new enclosed sequences not described elsewhere are part of the list:

- 1F14F SQUARED WC
= public toilet
- 1F18E NEGATIVE SQUARED AB
= blood type AB
- 1F17F NEGATIVE SQUARED WC
= public toilet

Editorial comments

E1. Amendment text.

For “Page 2, Clause 3 Normative references”, read “Page 2, Clause 3, Normative references”.

For “Page 14, Sub-clause 20.3 Format characters”, read “Page 14, Sub-clause, 20.3 Format characters”.

For “Page 25, Clause 29 Named UCS Sequence Identifiers”, read “Page 25, Clause 29, Named UCS Sequence Identifiers”.

For “Page 30-1348 Clause 34, Code Tables and list of character names”, read “Page 30-1348, Clause 34, Code Tables and list of character names”.

For “Page 1351, annex A.1”, read “Page 1351, Annex A.1”.

For “Page 1357, Annex A.6 Unicode Collections”, read “Page 1357, Annex A.6, Unicode Collections”.

For “Page 1358, Annex B List of combining characters”, read “Page 1358, Annex B, List of combining characters”.

Partially accepted

The “Page, Clause name’ format has been used for a long time in many existing amendments and does not need to be changed at this time. It will be used consistently (Ireland correctly spotted the inconsistency between the mention to ‘Page 30-1348’ and others). Furthermore ‘annex A.1’ will be replaced with ‘Annex A.1’”.

E2. Page 1358, Annex B, List of combining characters.

Check to remove a leading space before the word MYANMAR at 109A, 109B, 109C.

Accepted

E3. Page 7, Row 050: Cyrillic Supplement.

Ireland suggests that that the following notes be used for consistency:

Letters for Enets and Khanty

0510 CYRILLIC CAPITAL LETTER REVERSED ZE

..

Letters for Chukchi, Itelman, and Khanty

0512 CYRILLIC CAPITAL LETTER EL WITH HOOK

..

Letters for Aleut

051E CYRILLIC CAPITAL LETTER ALEUT KA

..

Letters for Abkhaz

0524 CYRILLIC CAPITAL LETTER PE WITH DESCENDER

0525 CYRILLIC SMALL LETTER PE WITH DESCENDER

- used in modern Abkhaz orthography

→ 04A7 cyrillic small letter pe with hook

Accepted in principle

The editor will work with Ireland and the Unicode editor to come up with a mutually acceptable content for these editorial annotations.

E4. Page 14, Row 090: Devanagari.

Ireland suggests that that the glyph for 0955 be centred like 0945.

Accepted

Pending a new font is provided to the editor.

E5. Page 20, Row 0F0: Tibetan.

Ireland suggests that that the following notes be used:

0FCC TIBETAN SYMBOL NOR BU BZHI -KHYIL

• the quadruple body symbol, a form of the svastika

→ 0FD6 left-facing svasti sign

Accepted in principle

See disposition of comment E.1 from UK.

E6. Page 55, Row 260: Miscellaneous Symbols.

Ireland suggests that the glyph for 26DF be made larger. We also request the following annotation:

26DF BLACK TRUCK

= black lorry

Accepted

See disposition of comment E.3 from USA.

E7. Page 72, Row A4D: Lisu.

Ireland recommends the replacement of the font used for Lisu with the table shown below. This font was prepared by the original submitter of the Lisu proposal, David Morse, and better reflects community glyph preferences.

Accepted

Pending a new font is provided to the editor.

E8. Page 100, Row 1F20: Enclosed Ideographic Supplement.

Ireland suggests that the annotation to 1F243 be changed to “safe” and that the annotation to 1F248 be changed to “steal”.

Accepted

See disposition of comment E.31 from UK.

India (Positive)

(Comment received from India NB outside of the ballot document SC2 N4065 in WG2 N3599.)

T1. Javanese

With regard to ISO/IEC 10646:2003/FPDAmd 6 Information technology - Universal Multiple-Octet Coded Character Set (UCS) - AMENDMENT 6: Javanese, Lisu, Meetei Mayek, Samaritan, and other characters, we received a small correction from our expert on Javanese character as follows :

On page 74, Character A9C0

Written : JAVANESE PANGKONG

Should be : JAVANESE PANGKON

Proposed acceptance

Germany, Positive with comments

Technical comments

T.1 Vedic Extensions – renaming for 1DC4

Germany supports the U.S. Comment expressed in the INCITS/L2/09-082 from 2009-02-6 (See comment T.1 from USA)

WG2 discussion

Similar comment from Ireland T.4. See disposition of comment T.1 from USA.

T.2 Enclosed Alphanumeric Supplement – renaming of some characters

Germany supports the U.S. Comment expressed in the INCITS/L2/09-082 from 2009-02-6 a) (See comment T.2 a) from USA)

Propose acceptance

Also supported by comment T.2a) from USA and T.9.a from Ireland.

b) (See comment T.2 b) from USA)

Propose acceptance

Also supported by comment T.2b) from USA.

T.3 Enclosed Alphanumeric Supplement – addition of WZ

Germany supports the U.S. Comment expressed in the INCITS/L2/09-082 from 2009-02-6 (See comment T.3 from USA)

Propose acceptance

Also supported by comment T.3 from USA.

T4. Unified Canadian Aboriginal Syllabics Extended block - reordering

Germany supports the U.S. Comment expressed in the INCITS/L2/09-082 from 2009-02-6 (See comment T.4 from USA)

Propose acceptance

Also supported by comment T.4 from USA and T.3 from Ireland.

T5. Kaithi –addition of 2 characters

Germany supports the U.S. Comment expressed in the INCITS/L2/09-082 from 2009-02-6 (See comment T.5 from USA)

Propose acceptance

Also supported by comment T.5 from USA and T.8 from Ireland.

T6. Miscellaneous Symbols block – move characters

Germany requests the following character moves, to accommodate the more iconic symbols placed in the Miscellaneous Symbols block and to put the other symbols in blocks which already contain similar symbols:

269E THREE LINES CONVERGING RIGHT,

269F THREE LINES CONVERGING LEFT:

into the Miscellaneous Technical block, at U+23EE/U+23EF or other positions in that block

26CE HEAVY EXCLAMATION MARK SYMBOL:

into the Dingbats block, at U+2700

26E2 HEAVY LARGE CIRCLE,

26E4 HEAVY OVAL WITH OVAL INSIDE,

26E5 HEAVY CIRCLE WITH CIRCLE INSIDE,

26E6 HEAVY CIRCLE,

26E7 HEAVY CIRCLED SALTIRE:

into the Miscellaneous Symbols and Arrows block, at U+2B55...U+2B59, to get these pure geometrical shapes together with other geometrical shapes.

To have the ARIB symbols in the Miscellaneous Symbols block in a contiguous area, then shift the remaining symbols:

26CF...26E1 back one position,

26E3 back two positions,

26E8...26FF back six positions.

This will leave two contiguous gaps in the Miscellaneous Symbols block: 269E...269F and 26F1...26FF.

WG2 decision

The editor has no preference. Either way can be accommodated.

T7. Miscellaneous Symbols block and Enclosed Alphanumeric Supplement block – glyph change

Germany requests two glyph changes, which both yield glyphs more similar to the original ARIB symbols. See p.58 and p.62 on: http://www.arib.or.jp/english/html/overview/doc/6-STD-B24v5_1-1p3-E1.pdf

26F0 MOUNTAIN (according to T6: 26EA MOUNTAIN): The top should be rounded. This also makes the symbol different from U+25B2 BLACK UP-POINTING TRIANGLE.

1F12D CIRCLED CD: The outline shall be an exact circle instead of an oval.

Accepted in principle

Pending access to improved glyphs. See also comment JPE7 from Japan.

Editorial Comments:

E8. Devanagari Extended – add annotation for A8FA

Germany supports the U.S. Comment expressed in the INCITS/L2/09-082 from 2009-02-6 (See comment E.1 from USA)

Accepted

E9. Old South Arabian – add annotation

Germany supports the U.S. Comment expressed in the INCITS/L2/09-082 from 2009-02-6 (See comment E.2 from USA)

WG2 discussion

Although Editorial comments are non technical by nature, these are cross referenced by technical comment T.7 from Ireland which requires a WG2 discussion, including name changes.

E10. Miscellaneous Symbols – improve black truck picture

Germany supports the U.S. Comment expressed in the INCITS/L2/09-082 from 2009-02-6 (See comment E.3 from USA)

Accepted