

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

Doc. Type: Disposition of comments

Title: Disposition of comments on SC2 N 3584 (PDAM text for Amendment 2 to ISO/IEC 10646-1:2000)

Source: Michel Suignard (project editor)

Project: JTC1 02.18.01

Status: For review by WG2

Date: 2001-05-21

Distribution: WG2

Reference: SC2 N3596, N3602, N3604

Medium: Paper

Comments were received from Canada, Germany, Ireland, Japan, DPRK, ROK, Sweden and USA. The following document is the draft disposition of those comments. The disposition is organized per country.

Note - The full content of the ballot comments (minus some character glyphs) 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 italic serif*.

As of result of the disposition of comments, Ireland and the USA changed their votes from No to Yes.

Canada: Yes with comments:

Technical comment:

Comment ii – Specify a Unicode 4.0 collection.

A new fixed collection identifier is requested to identify the repertoire of 10646-1 when its Amd.2 to Part 1 and Amd.1 to Part 2 will be finalized -- this will enumerate and identify the fixed repertoire of Unicode Version 4.0, similar to the fixed collection identifiers for the previous Unicode versions.

Accepted

A new Unicode collection (#305) will be described in a new sub-clause (A.5.3) covering these 2 amendments. This obviously exacerbates the synchronization issue between Part 1 and Part 2 .

Editorial Comments:

Comment i:

On page 3, changes marked for Page 20-303, item 1, second paragraph:

should be changed to "These tables contain new characters ..."
instead of "These tables contained new characters ..."

Accepted

Comment iii:

Change the annotation in the name from Rupal to Rupai (the last letter is lower case i instead of a lower case l or number 1), for 0B9F TAMIL RUPEE SIGN (Rupal).

Accepted

Germany: Yes with comments:

Technical comments:

Comment 1: Page v, Foreword:

Add a note to identify which parts of the standard are only available in machine readable form, preferably by pointing to the sections in which they are linked or, alternatively, by listing the relevant files here directly.

Furthermore, check the change against ISO's stated intention to improve the involvement of industrially less advanced countries in standardization. Experts from these countries may not have the means to access machine readable information. In this case, it might be necessary for ISO Central Secretariat to offer a version with all parts printed out.

Follow any advice from ISO Central Secretariat on this issue.

Partially accepted

The Foreword can identify explicitly by using filenames which parts of the standard are intently mostly as machine readable files. Because these materials are still printable (although not page formatted to preserve their machine readable feature) this does not constitute a barrier to access to industrially less advanced countries.

Recently, this standard has been distributed in a CD-ROM media form which contains all files which can be printed as desired by the recipients. CD-ROMs are much cheaper to reproduce than large documents and can be considered improvements to access to the standard.

The foreword notice will be changed as follows:

The standard contains material which may be only available to users who obtain their copy in a machine-readable format. That material consists of the following printable files:

CJKC0SR.txt
CJKUA_SR.txt
[...]

Comment 2: Item 9: Character U267F:

Remove this character from the document in order to take into account mail SC2.616 from 2002-02-20

It is desirable to encode this character. It is used frequently in many kinds of publications including many school books.

As suggested in the aforementioned mail, this encoding may be impossible due to the legal / trade mark considerations mentioned by Der Grüne Punkt - Duales System Deutschland AG. However, given the quasi-monopolist and state-backed position that this company still enjoys in Germany and given that the symbol is used far beyond German borders this may be a borderline case.

The company's claim that the symbol "has no meaning as a recycling symbol or a meaning as environmental friendly" is hardly in line with their own marketing position, even less with popular perception.

Therefore remove this character from this proposal until independent legal advice can be obtained.

If possible to maintain, rename to GREEN DOT (recyclable packaging)

Accepted

The character will be removed (removal also requested by the US comment T.9 and Ireland T.6).

Ireland: Negative

Technical comments

T1. Page 888, Annex B. As per the changes proposed in WG2 N2410, replace the entire list of Limbu entries with the following:

1920 LIMBU VOWEL SIGN A
1921 LIMBU VOWEL SIGN I
1922 LIMBU VOWEL SIGN U
1923 LIMBU VOWEL SIGN EE
1924 LIMBU VOWEL SIGN AI
1925 LIMBU VOWEL SIGN OO
1926 LIMBU VOWEL SIGN AU
1927 LIMBU VOWEL SIGN E
1928 LIMBU VOWEL SIGN O
1929 LIMBU SUBJOINED LETTER YA
192A LIMBU SUBJOINED LETTER RA
192B LIMBU SUBJOINED LETTER WA
1930 LIMBU SMALL LETTER KA
1931 LIMBU SMALL LETTER NGA
1932 LIMBU SMALL LETTER ANUSVARA
1933 LIMBU SMALL LETTER TA
1934 LIMBU SMALL LETTER NA
1935 LIMBU SMALL LETTER PA
1936 LIMBU SMALL LETTER MA
1937 LIMBU SMALL LETTER RA
1938 LIMBU SMALL LETTER LA
1939 LIMBU SIGN MUKPHRENG
193A LIMBU SIGN KEMPHRENG
193B LIMBU SIGN SA-I

Accepted

Note that some definitions are the same as the ones used in the PDAM2 text and this is consistent with the US comment T.1)

T2. Table 54 - Row 19: Limbu.

Replace the table with the table and names list on pages 5 and 6 below.

Accepted

Also requested by US comment T.1

T3. 23CF WHITE UP-POINTING TRIANGLE WITH CAUTION MARK and 23D0 HIGH VOLTAGE CAUTION MARK:

Neither of these characters is a technical symbol. These should be moved to 26A0 and 26A1.

Accepted

Also requested by US comment T.10 (which also asks for a name change)

T4. 2614 SHOWER WEATHER.

This is a glyph variant of 2602 UMBRELLA. There are many kinds of weather: showery weather, blustery weather, windy weather. Shall we also encode an inside-out umbrella to indicate gales? If accepted, this character should be renamed UMBRELLA WITH RAINDROPS and the glyph harmonized with 2602 (the glyph proposed for 2614 is better than that of 2602).

Accepted

Also requested by US comment T.10

T5. 2618 CIRCLED UPWARD INDICATION.

This is a glyph variant of 261D WHITE UP POINTING INDEX. The circled variant could be encoded using 20DD COMBINING ENCLOSING CIRCLE. If it must be encoded, its name should be CIRCLED WHITE UP POINTING INDEX, and its glyph should make use of the existing Zapf index at 261D.

Accepted in principle

The character was removed as result of the disposition of the US comment T.10. The proposed character is considered a glyph variant of the character 261D WHITE UP POINTING INDEX.

T6. 267F RECYCLABLE PACKAGING (Green Dot).

This is not a generic recyclable packaging symbol. It is a registered trademark of PRO Europe and its constituent bodies, whom we have contacted and who have indicated that they do *not* want this logo to be encoded in the UCS. This character must be removed from the PDAM.

Accepted

Similar to German 2nd comment and US comment T.9.

T7. 267E DO NOT LITTER SIGN.

Ireland would like to reiterate our support for this character, which was first proposed in N2310 (2000-12-22) and N2342 (2001-04-02). Like the UNIVERSAL RECYCLING SYMBOL, it is found in many generic contexts in printing, educational materials and packaging. We would, however, like to have its name changed to DO NOT LITTER SYMBOL. Among its aliases is the “Pitch-In Symbol”, and it seems that SYMBOL is a better term than SIGN for this, as for the other recycling symbols.

Withdrawn

T8. 2700 LEFTWARD SCISSORS.

This is a glyph variant of 2702 BLACK SCISSORS. Note that no RIGHTWARD SCISSORS (to be used by left-handers was proposed). Rotated versions of symbols are not generally encoded in the UCS, apart from arrows and some of the Zapf pointing fingers. This character should be deleted from the PDAM.

Accepted

The US comment T.10 is also requesting its removal

T9. 31D0-31D4

31D0 VULGAR FRACTION ONE HALF WITH HORIZONTAL BAR,
31D1 VULGAR FRACTION ONE THIRD WITH HORIZONTAL BAR,
31D2 VULGAR FRACTION TWO THIRDS WITH HORIZONTAL BAR,
31D3 VULGAR FRACTION ONE QUARTER WITH HORIZONTAL BAR,
31D4 VULGAR FRACTION THREE QUARTERS WITH HORIZONTAL BAR

These are glyph variants of

00BD VULGAR FRACTION ONE HALF
2153 VULGAR FRACTION ONE THIRD
2154 VULGAR FRACTION TWO THIRDS
00BC VULGAR FRACTION ONE QUARTER
00BE VULGAR FRACTION THREE QUARTERS

There is no justification whatsoever to make an artificial coding distinction between diagonal and horizontal fraction slashes. These characters should be deleted from the PDAM.

Accepted

The US comment T.9 is also requesting their removal

T10. 2615 TEA SYMBOL.

Change this name to HOT BEVERAGE; most characters in this block do not have the word SYMBOL in their names, and this is a picture of a hot beverage, not an abstract symbol like the RECYCLING SYMBOLS. We believe that a more generic less 3D cup such as those shown in N2417 might be better for the glyph.

Accepted

The US comment T.10 requested a similar change to 'HOT BEVERAGE SYMBOL' (note the added symbol). The consolidated disposition is to accept the name changed proposed by Ireland: 2615 HOT BEVERAGE.

T11. General comment. We don't much like all the pre-composed squared Latin abbreviations and would like to see fewer of them, and would like to see them filling holes in the existing CJK blocks if possible.

Accepted in principle

The US comment T.10 is requesting to move most of them to an existing CJK blocks and the last to the LetterLike Symbol block)

T12. FE47-FE48

FE47 PRESENTATION FORM FOR VERTICAL LEFT SQUARE BRACKET and
FE48 PRESENTATION FORM FOR VERTICAL RIGHT SQUARE BRACKET

and are duplicates of

23B4TOP SQUARE BRACKET and

23B5 BOTTOM SQUARE BRACKET.

Delete them.

Not accepted

The characters look similar but they are a different set of compatibility characters used in vertical forms of East Asian fonts.

T13. Table 101 - Row 31 Supplemental CJK Compatibility.

Why are these three columns wide? We would prefer to limit CJK Compatibility characters. It is 2002, now, not 1993.

Accepted in principle

The block has been removed as a result of disposing of other comments.

T14. Specification for the Modern European Scripts collection.

The MES-3A and MES-3B collections were removed by WG2 from PDAM 1 to ISO/IEC 10646-1:2000 pending clarification from CEN TC304 as to its satisfaction with them. CEN/TC304 approved the following in its Plenary on 3 December 2001 in Brussels:

Resolution # 9 MES-3 in 10646

CEN/TC304 resolves to suggest to SC2/WG2 that MES-3A and MES-3B not be incorporated in the Annex of 10646 as such. Instead, if incorporated they should be replaced by just one open collection, the "Modern European Scripts" (MES-Open), that reflects the scope of the CEN WS MES that produced MES-3A and MES-3B as part of the CEN Workshop Agreement 13873.

Accordingly, Ireland proposes the following collection be added to PDAM 2 of ISO/IEC 10646-1. It includes the new collection 93 CYRILLIC SUPPLEMENT which did not exist in 2000 when CWA 13873 was drawn up. In addition, to resolve conflict and uncertainty regarding the bidirectional nature of the characters within the existing collection 63 ALPHABETIC PRESENTATION FORMS, we reiterate the request that was made when all three MES collections were proposed by Ireland and Finland, namely, to add two new collections so that collection-based subsets can be conveniently made. (This idea was supported by TC304 though it did not appear in the minutes.)

xx LTR ALPHABETIC PRESENTATION FORMS FB00–FB17

yy RTL ALPHABETIC PRESENTATION FORMS FB18–FB4F

Collection Name: Modern European Scripts

Type of Collection (Fixed or Non-fixed): NON-FIXED

The following UCS collections from annex A of ISO/IEC 10646-1 comprise the Modern European Scripts collection:

No.	Collection name	hex range
1	BASIC LATIN	0020–007E
2	LATIN-1 SUPPLEMENT	00A0–00FF
3	LATIN EXTENDED-A	0100–017F
4	LATIN EXTENDED-B	0180–024F
5	IPA EXTENSIONS	0250–02AF
6	SPACING MODIFIER LETTERS	02B0–02FF
7	COMBINING DIACRITICAL MARKS	0300–036F
8	BASIC GREEK	0370–03CF
9	GREEK SYMBOLS AND COPTIC	03D0–03FF
10	CYRILLIC	0400–04FF
92	CYRILLIC SUPPLEMENT	0500–052F
11	ARMENIAN	0530–058F
27	BASIC GEORGIAN	10D0–10FF
30	LATIN EXTENDED ADDITIONAL	1E00–1EFF
31	GREEK EXTENDED	1F00–1FFF
32	GENERAL PUNCTUATION	2000–206F
33	SUPERSCRIPTS AND SUBSCRIPTS	2070–209F
34	CURRENCY SYMBOLS	20A0–20CF
35	COMBINING DIACRITICAL MARKS FOR SYMBOLS	20D0–20FF
36	LETTERLIKE SYMBOLS	2100–214F
37	NUMBER FORMS	2150–218F
38	ARROWS	2190–21FF
39	MATHEMATICAL OPERATORS	2200–22FF
40	MISCELLANEOUS TECHNICAL	2300–23FF
42	OPTICAL CHARACTER RECOGNITION	2440–245F
44	BOX DRAWING	2500–257F
45	BLOCK ELEMENTS	2580–259F
46	GEOMETRIC SHAPES	25A0–25FF
47	MISCELLANEOUS SYMBOLS	2600–26FF
xx	LTR ALPHABETIC PRESENTATION FORMS	FB00–FB17
65	COMBINING HALF MARKS	FE20–FE2F
70	SPECIALS	FFF0–FFFF

Accepted

The name of the collection is MODERN EUROPEAN SCRIPTS. The content of the collections xx and yy (values to be determined by the editor) is modified as follows:

xx LTR ALPHABETIC PRESENTATION FORMS FB00-FB1C

yy RTL ALPHABETIC PRESENTATION FORMS FB1D-FB4F

T15. Addition of Symbols.

Ireland recommends the addition of a number of symbols to the PDAM.

T15-1. Ireland proposes the addition of two Greek letters for Bactrian, as given in N2411,

GREEK CAPITAL LETTER BACTRIAN SHO at 03F7 and

GREEK SMALL LETTER BACTRIAN SHO at 03F8.

Accepted

T15-2. audio-visual media symbols

Ireland proposes the addition of seven user-interface symbols for removable and audio-visual media, as discussed in N2415. These are:

23CF EJECT SYMBOL

23D0 PAUSE

23D1 FAST FORWARD

23D2 FAST REVERSE

23D3 GO TO BEGINNING

23D4 GO TO END

2614 MEDIUM CIRCLE (start recording)

Partially accepted

The character EJECT SYMBOL will be added (also requested by the US comment T7.b). The glyph will be shown in its filled version (not white). The other proposed characters are postponed for further study.

T15-3. map and guidebook symbols

Ireland proposes the addition of nine map and guidebooks symbols related to the WHITE FLAG and BLACK FLAG in this PDAM, as discussed in N2415. These are:

2692 CROSSED WHITE AND BLACK FLAGS

2693 BATTLEFIELD MARKER

2694 MONASTERY

2695 CATHEDRAL

2696 CASTLE

2697 BISHOPRIC

2698 ARCHBISHOPRIC

2699 HEBREW TEMPLE

269A MOSQUE

Withdrawn

(The actual referenced document should have been WG2 N 2486)

T15-4. Khmer

Ireland is aware of a set of Khmer characters proposed to be added by the UTC. While we await a WG2 document describing the functions of these characters, we would like to make the following comment on the names proposed by the UTC.

a) Firstly, the characters for Krung must have the word KHMER in the first position as the script identifier. It would be good to know what the native name for these is and to either use that or translate it.

Moot

The characters for Krung were not added as a result of disposing the US comment T8.

b) Secondly, the glyphs for the lunar calendar symbols are numeric, and the names are composed of digit names (for the numbers 1 to 15) and other elements.

In the UCS, it is longstanding practice that such names are translated into English, and we strongly suggest that this practice be extended to these character names for greater transparency to implementers of the Khmer encoding.

Withdrawn

Editorial comments

E1. Page 881, annex A.1.

Change "annex" to "Annex" in the title of this item.

Accepted

E2. Table 25 - Row 0B: Tamil.

The newly added characters are not in the same font as the characters from the rest of the table. This should be corrected.

Accepted

As of result of the disposition of the comments Ireland changes its vote to Yes.

Japan: Yes with comments:

Technical comments:

1. Limbu

NB of Nepal indicates to NB of Japan that Nepal has three counter proposals. Change of character name for some of the characters, Addition of 5 Limbu characters and Reservation of 6 code positions for future extension. None of them are requesting existing code position change. Therefore, NB of Japan requests to reflect the character name change request on this PDAM. The NB comment of Nepal will be submitted in separated paper.

Accepted in principle

The comments from Nepal are included in document WG2 N2428-1:5. Comments made by the US (comment T1) and Ireland (T1) are addressing the same subject. Disposing of these comments is also covering in large part the requests from Nepal. Nepal is also requesting to add symbols which are already contained in the General Punctuation block and therefore should not be re-encoded. The only missing symbol (LIMBU PUNCTUATION HYPHEN) looks like a double hyphen and its usage need to be evaluated against other scripts before being encoded.

Korea, Democratic People's Republic (DPRK), Yes with comments

Technical Comments

T1. CJK Mapping table.

In the file 'CJKUA_SR.txt', map the code of <KP1-61A9> to <U06B76>.

Accepted

Note that this is a new entry; KP1-61A9 didn't exist in the current CJKUA_SR.txt file. The WG2 reference document is N2426.

T2. CJK Compatibility Ideographs extension

Add of 144 Compatibility ideographs into the CJK Compatibility Ideographs.

[Note]

M41.10 WG2 invites DPR of Korea to examine their proposal for 160 CJK compatibility characters towards reducing the total number to maximum of 151 with the goal of being able to fit them into the available space in the BMP. WG2 further invites DPR of Korea to input this revised proposal as part of their ballot comments to Amendment 2 of IO/IEC 10646-1:2000.

According to this resolution, we reduced our compatibility ideographs for BMP to 144 characters. We already sent the document on 144 compatibility ideographs to Mr. Suignard, Chief editor of WG2 on 6th Dec, 2001 by E-mail. We also sent the TrueType Font and code mapping table of the 144 Compatibility ideographs to Mr. Asumusu [sic], representative of Unicode Consortium.

Accepted in principle

Following further discussion, it was agreed to reduce the addition to 122 new CJK Compatibility Ideographs at positions FA70-FAE9 and map the rest to existing CJK Compatibility Ideographs in plane 0 and 2. The reference document is WG2 N2493.

Editorial comments

E1. Table 103.

Change glyph of <U321E> in Table 103-Row 32 of page 26 in the ISO/IEC 10646-1:2000/PDAM2 into (오 후).

Accepted

The parentheses should not be 'white' (font production issue)

E2. Replace KPS by KP in clause 27.1

In 27.1 of page 1 in the ISO/IEC 10646-1:2000/PDAM2, replace 'Insert the new Hanja KPS source references after the Hanja K sources as follows:

Hanja KPS sources are

KP0 KPS 9566-97

KP1 KPS 10721-2000'

By

'Insert the new Hanja KP source references after the Hanja K sources as follows:

Hanja KPS sources are

KP0 KPS 9566-97

KP1 KPS 10721-2000'

Accepted

Korea, Republic of (ROK), Negative

Technical Comments

T1. Symbol name issue.

Names for the following 4 symbols:

U+0321D [PARENTHEZIZED KOREAN CHARACTER] OJEON

U+0321E [PARENTHEZIZED KOREAN CHARACTER] OHU

U+0327C [CIRCLED KOREAN CHARACTER] CHAMKO

U+0327D [CIRCLED KOREAN CHARACTER] JUEUI

We used 11941:1996 for naming Korean Hangul letters and syllables. However, the transliteration used for naming the following characters in PDAM2 is different from TS 11941.

Considering that, in the future, we may add more Korean Hangul letters in addition to the above characters, we suggest that we set up a transliteration rule for naming Korean Hangul letters and syllables and adhere to that rule.

Not accepted

We could not get new names that would be acceptable to the rest of the committee.

Sweden: Yes with comments:

Technical comments:

SE 1. MES

CEN/TC304 resolves to suggest to SC2/WG2 that MES- 3A and MES-3B not be incorporated in the Annex of 10646 as such. Instead if incorporated they should be replaced by just one open collection, the "Modern European Scripts" (MES-Open), that reflects the scope of the CEN WS MES that produced MES-3A and MES-3B as part of the CEN Workshop Agreement 13873:2000.

Accepted

As a result of disposing of Irish comment T.14.

SE 2.

page 1: "3.x" --> "4.0"?

Accepted

SE 3. (Data file format)

page 2: A semicolon-separated format for the data file would be preferable, and would avoid the byte counting oddity. Say explicitly that the encoding of the data file is ISO/IEC 646 IRV with LINE FEED as end of line mark.

Partially accepted

A similar comment was made during the creation of Part which introduced the concept of linked data files. A similar comment was made for Part 2 when the files were introduced. These data files are mechanically generated by IRG and it would be very unwise to manually change the format of these as it could introduce errors. The format is easy to import by database tools as it stands.

The second part of the comment is already covered by the mention to ISO/IEC 646-IRV characters (section 27.2 in page 1). The only part that will be added concerns 'LINE FEED as end of mark'. Note however that all clauses that use the same languages in both parts need to be updated accordingly. (Another argument for merging part 1 and 2)

SE 4.

page 2: “code value” --> “code position” (use the same terminology throughout 10646)

Accepted

All other occurrences of the same text will be editorially fixed in both parts of the standard.

SE 5.

U+0BF3 – U+0BFA: for these characters the representative glyphs have a different font design compared to what is used for the related Tamil characters. The font designs for the example glyphs should be the same for all characters in the same script.

Accepted in principle

Based on the editor receiving a new font.

SE 6.

U+31E0 – U+31E4: The characters here are already allocated (though the sample glyphs given then use a slanted design. If a plain text level distinction need be made, that should be done using variation selectors. Note that then the variation selector approach must work for all vulgar fractions, including those composed using FRACTION SLASH (U+2044).

Accepted in principle

The removal of these characters is also requested by Ireland (comment T.9) and the US (comment T.9).

SE 7.

The characters proposed for U+31C0 – U+31C5 are further abbreviations in “square block” form. Given that those abbreviations (already allocated) are far from complete, and there should be no effort to complete them, instead such “squared” abbreviations should be recommended against. That implies that no new ones should be allocated.

The same holds for U+321D--U+321E and U+327C--U+327D, and U+3377—U+337A.

Withdrawn

Also mentioned in US comment T.10.

USA: No with comments:

Technical comments:

T.1 Limbu

After further study on the Limbu script described by WG2 N2410, The US National body requests the following changes to the proposed Limbu script.

The following characters should be added:

1909 LIMBU LETTER JHA
190A LIMBU LETTER YAN
191A LIMBU LETTER SSA
1932 LIMBU SMALL LETTER ANUSVARA
1940 LIMBU SIGN LOO
1944 LIMBU EXCLAMATION MARK

The Limbu vowel signs should be re-arranged as follows (former name and location in parenthesis)

1920 LIMBU VOWEL SIGN A (1920 LIMBU VOWEL SIGN A)
1921 LIMBU VOWEL SIGN I (1922 LIMBU VOWEL SIGN I)
1922 LIMBU VOWEL SIGN U (1923 LIMBU VOWEL SIGN U)
1923 LIMBU VOWEL SIGN EE (1924 LIMBU VOWEL SIGN E)
1924 LIMBU VOWEL SIGN AI (1925 LIMBU VOWEL SIGN AI)
1925 LIMBU VOWEL SIGN OO (1926 LIMBU VOWEL SIGN O)
1926 LIMBU VOWEL SIGN AU (1927 LIMBU VOWEL SIGN AU)
1927 LIMBU VOWEL SIGN E (1929 LIMBU VOWEL SIGN EE)
1928 LIMBU VOWEL SIGN O (1928 LIMBU VOWEL SIGN OO)
1929 LIMBU SUBJOINED LETTER YA (192A LIMBU SUBJOINED LETTER Y)
192A LIMBU SUBJOINED LETTER RA (192B LIMBU SUBJOINED LETTER R)
192B LIMBU SUBJOINED LETTER WA (192C LIMBU SUBJOINED LETTER W)

The existing characters between 1930 and 1938 should be renamed as follows ('A' appended to their name):

1930 LIMBU SMALL LETTER KA
1931 LIMBU SMALL LETTER NGA
1933 LIMBU SMALL LETTER TA
1934 LIMBU SMALL LETTER NA
1935 LIMBU SMALL LETTER PA
1936 LIMBU SMALL LETTER MA
1937 LIMBU SMALL LETTER RA
1938 LIMBU SMALL LETTER LA

The character in code position 193F should be moved to 1945 (by similarity with the addition in position 1944) with the slightly modified name as follows:

1945 LIMBU QUESTION MARK

Accepted

This request is compatible with Irish comment T.1 and T.2

T.2 Normalization reference

At the last WG2 meeting in Singapore, it was proposed to add a normative reference to normalization into the standard (document WG2 N2398). The US national body is asking to add the text proposed by N2398 into this amendment. The additions and changes are as follows (quoted from N2398)

<<

In clause '3 Normative reference', add the following reference:

Unicode Standard Annex, UAX#15, Unicode Normalization Forms, Version 3.1., 2001-03-23.

Insert a new clause 24 between the current clause 23 and 24 (the following clauses are incremented by one):

24 Normalization forms

Normalization forms are the mechanisms allowing the selection of a unique coded representation among alternative, but equivalent coded text representations of the same text. Normalization forms for use with ISO/IEC 10646 are specified in the Unicode Standard UAX#15.

NOTE 1 – By definition, the result of applying any of these normalization forms is stable over time. It means that a normalized representation of text remains normalized even when the standard is amended.

NOTE 2 – Some normalizations forms favor composite sequences over shorter representations of text, other favor the shorter representations. The backward compatibility requirement is provided by establishing ISO/IEC 10646-1:2000 and ISO/IEC 10646-2:2001 as the reference versions for the definition of the shorter representation of text.

In the current clause 24 (which becomes 25), insert a new clause 25.3 (after the current clause 24.2):

25.3 Alternate coded representations

Alternate coded representations of text are generated by using multiple combining characters in different orders, or using various equivalent combinations of characters and composite sequences. These alternate coded representations result in multiple representation of the same text. Normalizing these coded representations creates a unique representation.

NOTE – For example, in implementation level 3 the French word "là" may be represented by the characters LATIN SMALL LETTER L followed by LATIN SMALL LETTER A WITH GRAVE, or may be represented by the characters LATIN SMALL LETTER L followed by LATIN SMALL LETTER A followed by COMBINING GRAVE ACCENT. When the normalizations forms are applied on those alternate coded representations, only one representation remains. The form of the remaining representation depends on the normalization form used.

In the current clause 24.3 (new 25.4), remove the Note (replaced by note mentioned above)

>>

Accepted

This was requested by resolution M41.3 in the WG2 Singapore meeting (October 2001) and was not included in the PDAM2 text.

T.3 Arabic characters addition

The US I asking for the addition for the following Arabic symbols with the proposed code location and name (reference documents L2/01-427 and L2/02-061 with WG2 documents to be provided):

0600 ARABIC NUMBER SIGN
 0601 ARABIC YEAR SIGN
 0602 ARABIC FOOTNOTE MARKER
 060D ARABIC POETIC VERSE SIGN
 060E ARABIC DATE SEPARATOR
 0610 ARABIC SIGN SALLALLAHOU ALAYHE WASALLAM
 0611 ARABIC SIGN ALAYHE ASSALAM
 0612 ARABIC SIGN RAHMATULLAH ALAYHE
 0613 ARABIC SIGN RADI ALLAHOU ANHU
 0614 ARABIC SIGN NOM DE PLUME
 0656 ARABIC SUBSCRIPT ALEF
 0657 ARABIC TURNED DAMMA
 0658 ARABIC BREVE
 06EE ARABIC LETTER DAL WITH INVERTED V
 06EF ARABIC LETTER REH WITH INVERTED V
 06FF ARABIC LETTER HEH WITH INVERTED V

Three of the new characters (0600-602) should be added to a new category of format characters (see T.4). The characters 0610-0614, 0656-0658 needs to be added to annex B.1.

Accepted in principle

The reference documents are WG2 N2481-2483 and N2487. Some names were changed as a result of expert discussions with the following result (name changes underlined):

0600 ARABIC NUMBER SIGN
~~0601 ARABIC SIGN SANAH~~
 0602 ARABIC FOOTNOTE MARKER
 060D ARABIC DATE SEPARATOR
 060E ARABIC POETIC VERSE SIGN
 0610 ARABIC SIGN SALLALLAHOU ALAYHE WASALLAM
 0611 ARABIC SIGN ALAYHE ASSALAM
 0612 ARABIC SIGN RAHMATULLAH ALAYHE
 0613 ARABIC SIGN RADI ALLAHOU ANHU
~~0614 ARABIC SIGN TAKHALLUS~~
 0656 ARABIC SUBSCRIPT ALEF
~~0657 ARABIC INVERTED DAMMA~~
~~0658 ARABIC NOON-GHUNNA~~
 06EE ARABIC LETTER DAL WITH INVERTED V
 06EF ARABIC LETTER REH WITH INVERTED V
 06FF ARABIC LETTER HEH WITH INVERTED V

T.4 New format characters (general format characters and head marker)

There should be a brief description of the Combining Grapheme Joiner introduced by 10646-1:2001 Amd 1. This could be satisfied by adding a new entry in sub-clause F.1.1 as follows:

COMBINING GRAPHEME JOINER (034F): The Combining Grapheme Joiner is used to indicate that adjacent characters belong to the same grapheme cluster. Grapheme clusters are sequences of one or more coded characters that correspond to what users think of as characters. They include, but are not limited to, composite sequences such as (g + °), digraphs such as Slovak “ch”, or sequences with letter modifiers such as k^w.

In addition, there is a need to describe a type of alternate format character represented by the SYRIAC ABBREVIATION MARK (070F) as well as the ARABIC END OF AYAH (06DD), the latter currently classified as a combining character. The US national body is asking for the addition of a new clause in Annex F as follows:

F.5 Subtending format characters

The following characters are used to subtend a sequence of subsequent characters:

ARABIC NUMBER SIGN (0600)
 ARABIC YEAR SIGN (0601)
 ARABIC FOOTNOTE MARKER (0602)
 ARABIC END OF AYAH (06DD),
 SYRIAC ABBREVIATION MARK (070F)

The scope of these characters is the subsequent sequence of digits (plus certain other characters), with the exact specification as defined in Unicode 3.2 for ARABIC END OF AYAH.

In addition the ARABIC END OF AYAH (06DD) should be removed from Annex B – List of combining characters and its glyph should be updated in the chart to remove the dotted circle and use a dotted rectangular frame to be consistent with other format characters.

Accepted

T5. Korean CJK compatibility character mapping

The CJK COMPATIBILITY IDEOGRAPH F951 is mapped by the new clause 22.2 (Amd.1) to the CJK UNIFIED IDEOGRAPH 96FB, it should be mapped to 964B instead. The corresponding entry in the file CJKC0SR.txt needs to be updated.

Accepted

T6. Variation selectors

In sub-clause 24.5 (Amd.1), the following table row should be fixed as follows:

<2269, FE00>	GREATER-THAN BUT NOT EQUAL TO with vertical stroke
--------------	--

Accepted

T.7 Addition of symbols

a) The US National Body recommends adding the character NEGATIVE CIRCLED DIGIT ZERO to complete the series of negative circled digits U+2776-U+277F, U+24EB-24F4. This makes it similar to U+24EA, U+2460-U+2473. The proposed code point and name are as follows:

24FF NEGATIVE CIRCLED DIGIT ZERO

This makes the collection 43 ENCLOSED ALPHANUMERICS fixed.

Accepted

Proposal summary form WG2 N2441

b) The US also recommends adding the EJECT SYMBOL as a required user interface symbol for IT usage with code position and name as follows (document L2/01-414 to be submitted as a WG2 document):

23CF EJECT SYMBOL

Accepted

Ireland (comment T15-2) also asked for the addition of the EJECT SYMBOL at the same location. Proposal summary form WG2 N2432

T.8 Khmer

Following discussions with representatives of the Khmer Standardization national body, the US national body is asking for the addition of the following Khmer characters. An appropriate WG2 document will be provided.

17EA KRUNG VOICING OR PREGLOTTALIZATION SIGN
17EB KRUNG VOICING AND PREGLOTTALIZATION SIGN
17DD KHMER SIGN ATTHACAN
17F0 KHMER SYMBOL LEK ATTAK SON
17F1 KHMER SYMBOL LEK ATTAK MUOY
17F2 KHMER SYMBOL LEK ATTAK PII
17F3 KHMER SYMBOL LEK ATTAK BEI
17F4 KHMER SYMBOL LEK ATTAK BUON
17F5 KHMER SYMBOL LEK ATTAK PRAM
17F6 KHMER SYMBOL LEK ATTAK PRAM-MUOY
17F7 KHMER SYMBOL LEK ATTAK PRAM-PII
17F8 KHMER SYMBOL LEK ATTAK PRAM-BEI
17F9 KHMER SYMBOL LEK ATTAK PRAM-BUON
19E0 KHMER SYMBOL PATHAMASAT
19E1 KHMER SYMBOL MUOY KOET
19E2 KHMER SYMBOL PII KOET
19E3 KHMER SYMBOL BEI KOET
19E4 KHMER SYMBOL BUON KOET
19E5 KHMER SYMBOL PRAM KOET
19E6 KHMER SYMBOL PRAM-MUOY KOET
19E7 KHMER SYMBOL PRAM-PII KOET

19E8 KHMER SYMBOL PRAM-BEI KOET
 19E9 KHMER SYMBOL PRAM-BUON KOET
 19EA KHMER SYMBOL DAP KOET
 19EB KHMER SYMBOL DAP-MUOY KOET
 19EC KHMER SYMBOL DAP-PII KOET
 19ED KHMER SYMBOL DAP-BEI KOET
 19EE KHMER SYMBOL DAP-BUON KOET
 19EF KHMER SYMBOL DAP-PRAM KOET
 19F0 KHMER SYMBOL TUTEYASAT
 19F1 KHMER SYMBOL MUOY ROC
 19F2 KHMER SYMBOL PII ROC
 19F3 KHMER SYMBOL BEI BOC
 19F4 KHMER SYMBOL BUON ROC
 19F5 KHMER SYMBOL PRAM ROC
 19F6 KHMER SYMBOL PRAM-MUOY ROC
 19F7 KHMER SYMBOL PRAM-PII ROC
 19F8 KHMER SYMBOL PRAM-BEI ROC
 19F9 KHMER SYMBOL PRAM-BUON ROC
 19FA KHMER SYMBOL DAP ROC
 19FB KHMER SYMBOL DAP-MUOY ROC
 19FC KHMER SYMBOL DAP-PII ROC
 19FD KHMER SYMBOL DAP-BEI ROC
 19FE KHMER SYMBOL DAP-BUON ROC
 19FF KHMER SYMBOL DAP-PRAM ROC

Accepted in principle

After further discussion with Khmer experts it was agreed to postpone the encoding of the Krung characters (code position 17EA and 17EB). A new block will be created at the code range 19E0-19FF called KHMER SYMBOLS with the corresponding fixed collection.

T.9 Symbols to be removed

a) The characters 267E DO NOT LITTER SIGN and 267F RECYCLABLE PACKAGING must be removed. These characters have created many issues concerning usage and identity, so the US doesn't consider them mature for encoding.

Accepted

b) The character 31D0-31D4 (VULGAR FRACTION ... with horizontal bars) are duplicates of characters already encoded in the LATIN-1 SUPPLEMENT and NUMBER FORMS blocks and must not be encoded again as character. The existing characters have always been explicitly intended to unify both forms of the fraction.

Accepted

T.10 DPRK symbols

a) The name of the character 2614 SHOWER WEATHER should be changed to 'UMBRELLA WITH RAIN DROPS'.

In accordance with the glyph used for 2614, the graphic representation of 2602 UMBRELLA should be updated accordingly.

The name of the character 2615 TEA SYMBOL should be changed to 'HOT BEVERAGE SYMBOL'.

Accepted in principle

Similar to Irish comment T10. The resulting name is 'HOT BEVERAGE' (without the 'symbol').

b) The character 2618 CIRCLED UPWARD INDICATION seems to be a variation of 261D WHITE UP POINTING INDEX and may not be needed. At minimum, the name should be changed to reflect the similarity with 261D.

Accepted

The character is removed. Also requested by Ireland comment T.5

c) The character 2700 LEFTWARDS SCISSORS seems a variation of 2702 SCISSORS and should be removed accordingly.

Accepted

The character is removed. Also requested by Ireland comment T.7.

d) The characters a position 23CF WHITE UP-POINTING TRIANGLE WITH CAUTION MARK and 23D0 HIGH VOLTAGE CAUTION MARK should be moved and renamed as follows:

26A0 WARNING SIGN
26A1 HIGH VOLTAGE SIGN

Accepted

e) The block Supplemental Arrows-C (in row 2B) should be renamed ‘Miscellaneous Symbols and Arrows’.

Accepted

f) The first 6 characters of the newly proposed Supplemental CJK Compatibility should be moved either in holes of existing CJK blocks (for the first five characters: 31C0-31C4) or in the Letterlike Symbols block (for the fifth: 31C5) with the new location as follows:

32CC SQUARE HG
32CD SQUARE ERG
32CE SQUARE EV
32CF LIMITED LIABILITY SIGN
3250 PARTNERSHIP SIGN
213B FACSIMILE SIGN

Therefore, there is no need yet for the Supplemental CJK Compatibility block and it should be removed.

Accepted

Consequently, the Supplemental CJK Compatibility block is removed. Also mentioned in Swedish comment SE.7

T.11 Indic scripts character addition

The US is asking for the addition of the following characters with the suggested code position and names:

0904 DEVANAGARI SHORT LETTER A
09BD BENGALI SIGN AVAGRAHA
0A01 GURMUKHI SIGN ADAK BINDI
0A03 GURMUKHI SIGN VISARGA
0A8C GUJARATI LETTER VOCALIC L
0AE1 GUJARATI LETTER VOCALIC LL
0AE2 GUJARATI VOWEL SIGN VOCALIC L
0AE3 GUJARATI VOWEL SIGN VOCALIC LL
0AF1 GUJARATI RUPEE SIGN
0B35 ORIYA LETTER VA
0CBC KANNADA SIGN NUKTA
0CBD KANNADA SIGN AVAGRAHA

These additions are the results of several communications from the Indian government, in response to their review of ISO 10646 (reference L2/01-430R and L2/01-431R, WG2 document to be provided)

Accepted

The WG2 reference document is N2425.

As of result of the disposition of the comments the USA changes its vote to Yes.

[end]