INCITS/L2/09-304
Date: August 15, 2009
Title: Comments accompanying the U.S. negative vote on PDAM 8 to ISO/IEC 10646:2003 (SC2 N4078)
Source: INCITS/L2
Action: Forward to INCITS

## AMENDMENT 8

The U.S. National Body is voting No with comments on the following SC2 ballot: SC2 N4078:
Information technology -- Universal Multiple-Octet Coded Character Set (UCS) --
AMENDMENT 8: Additional symbols, Bamum supplement, CJK Unified Ideographs Extension
D, and other characters. If comments T.3.d, T.3.e., and T5 are accommodated, the U.S. will change its vote to Yes.

## Technical Comments:

## T.1. Malayalam

During the review of the ballot, the U.S. determined that one more character was required for Malayalam. As a result, the U.S. requests the following character be added to this amendment: U+0D4E MALAYALAM LETTER DOT REPH. The glyph, rationale, and properties are provided in N3676 (L2/09-245).

## T.2. Miscellaneous Symbols

While reviewing the 2600 block, which includes a new zodiacal symbol, OPHIUCHUS, U.S. noted an additional character was missing. As a result, the U.S. requests the addition of one character, U+26E2 ASTRONOMICAL SYMBOL FOR URANUS, with glyph and background as given in N3672 (L2/09-300).

## T.3. Miscellaneous Pictographic Symbols

The U.S. requests a series of changes to characters in Miscellaneous Pictographic Symbols.
a. Change the following three character names:

Change U+1F4AA FLEXED BICEP to: FLEXED BICEPS
(Rationale: biceps is the Oxford English Dictionary form for the singular)
Change U+1F69A TRUCK to: DELIVERY TRUCK
(Rationale: the name change serves to differentiate it from U+26DF BLACK TRUCK) Change U+1F3A5 MOVIES to: MOVIE CAMERA
(Rationale: the character represents a camera, not a projector; "movie camera" is unambiguous)
b. Change the annotation for U+1F3A5 MOVIE CAMERA (with the new name, as above) from
"film, movie projector" to "film, movies"
(Rationale: this better reflects its use in the Japanese sources)
c. Make three glyph changes:

U+1F526 BOOKMARK
(Rationale: make the glyph appear more like a bookmark, and not a luggage tag or a sales tag)
U+1F3A5 MOVIE CAMERA
(Rationale: see above)
U+1F420 CAMEL
(Rationale: the representative symbol should be a 2-hump camel which is reflected in the Japanese sources)
d. Add the following eleven characters to this amendment as listed below:

1F565 EMOJI COMPATIBILITY SYMBOL-6

* represents KDDI Emoji \#703
* represents Softbank Emoji \#430
= e-4ED
1F566 EMOJI COMPATIBILITY SYMBOL-7
* represents KDDI Emoji \#700
* represents Softbank Emoji \#425
= e-4E8
1F567 EMOJI COMPATIBILITY SYMBOL-8
* represents KDDI Emoji \#366
* represents Softbank Emoji \#428
= e-4EB
1F568 EMOJI COMPATIBILITY SYMBOL-9
* represents KDDI Emoji \#499
* represents Softbank Emoji \#424
= e-4E7
1F569 EMOJI COMPATIBILITY SYMBOL-10
* represents KDDI Emoji \#702
* represents Softbank Emoji \#427
= e-4EA
1F56A EMOJI COMPATIBILITY SYMBOL-11
* represents KDDI Emoji \#701
* represents Softbank Emoji \#426
= e-4E9
1F56B EMOJI COMPATIBILITY SYMBOL-12
* represents KDDI Emoji \#237
* represents Softbank Emoji \#422
= e-4E5
1F56C EMOJI COMPATIBILITY SYMBOL-13

```
    * represents KDDI Emoji #704
    * represents Softbank Emoji #431
    = e-4EE
1F56D EMOJI COMPATIBILITY SYMBOL-14
    * represents KDDI Emoji #367
    * represents Softbank Emoji #429
    = e-4EC
1F56E EMOJI COMPATIBILITY SYMBOL-15
    * represents KDDI Emoji #90
    * represents Softbank Emoji #423
    = e-4E6
1F56F EMOJI COMPATIBILITY SYMBOL-16
    * represents docomo Emoji #122
    = e-82B
```

The glyphs for the above eleven characters should be dashed boxes with text on two lines, "EC" on the top and the numbers ranging from $6 . .16$ on the bottom, similar to $\mathrm{U}+\mathrm{FFFC}$ and $\mathrm{U}+1680$.
e. The U.S. requests the inclusion of the emoji source data as documented in $\mathrm{L} 2 / 09-078=\mathrm{N} 3585$, as a linked data file named EmojiSources.txt, and addition of an appropriate clause in the standard that refers to the data file.

## T.4. Latin Extended D

The U.S. requests eight phonetic subscripts located at the codepoint locations U+A7F2-U+A7F9 in the Latin Extended D block be moved to the Superscripts and Subscripts block, where they will be included with other subscripts, and hence be more discoverable.

The characters are (with newly proposed codepoints):
2095 LATIN SUBSCRIPT SMALL LETTER H
2096 LATIN SUBSCRIPT SMALL LETTER K
2097 LATIN SUBSCRIPT SMALL LETTER L
2098 LATIN SUBSCRIPT SMALL LETTER M
2099 LATIN SUBSCRIPT SMALL LETTER N
209A LATIN SUBSCRIPT SMALL LETTER P
209B LATIN SUBSCRIPT SMALL LETTER S
209C LATIN SUBSCRIPT SMALL LETTER T

## T.5. Latin Extended D

The U.S. requests the removal of U+A78F LATIN LETTER MIDDLE DOT as we feel this character is unnecessary. Justification for this request is contained in N3678 (L2/09-278).

## T.6. Source Data for Ideographs

The U.S. strongly recommends that documentation of the source references for Ext. D be
provided. In order for Extension D to be included as part of an international standard, we feel that it is necessary to provide clear and complete source information in the form of a short summary of how to interpret the source data supplied in the Extension D document itself. This information is critical to properly interpret the characters in the Extension. For example, this summary could state that U-source data refer to the character's index within UTR \#45, with a URI indicating where UTR \#45 can be found.

