

INCITS/L2/10- 066

Date: February 5, 2010

Title: Comments accompanying the U.S. negative vote on FPDAM 8 to ISO/IEC 10646:2003 (SC2 N4123)

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 N4123: 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.1, T.2, T.5, and T.6.a are accommodated, the U.S. will change its vote to Yes.

Technical Comments:

T.1. Latin Extended D

The U.S. requests the removal of U+A78F LATIN LETTER MIDDLE DOT and reiterates that this character is unnecessary and is a damaging duplication for the standard. Justification for this request is contained in N3678 (L2/09-278). A viable alternative to encoding a separate letter middle dot, for the purposes cited by the original proposal, would be to use the already encoded modifier letter, U+02D1 MODIFIER LETTER HALF TRIANGULAR COLON.

T.2. Batak

The U.S. requests the removal of two Batak symbols:

1BFA BATAK SYMBOL BINDU GODANG

1BFB BATAK SYMBOL BINDU PINARJOLMA

Rationale: The evidence provided in N3320 does not demonstrate these marks to be characters, but to be graphic page elements that do not behave at all as characters. (See figures 11 and 13 in N3320 for clear examples showing that these are not characters.)

T.3. Alchemical Symbols

The U.S. requests three name changes in the Alchemical Symbols block, all involving a spelling change from "SULPHUR" to "SULFUR":

1F70D ALCHEMICAL SYMBOL FOR SULFUR

1F70E ALCHEMICAL SYMBOL FOR PHILOSOPHERS SULFUR

1F70F ALCHEMICAL SYMBOL FOR BLACK SULFUR

Rationale: The spelling with “F” is the one approved by the International Union of Pure and Applied Chemistry (1990) and the Royal Society of Chemistry, and is the consensus spelling used internationally.

T.4. Emoticons

a. The U.S. requests 1F62A NEUTRAL FACE be added to the Emoticons block, with the glyph, rationale, and properties as described in N3769 (L2/10-036).

b. The U.S. recommends the following 15 name changes suggested in N3711:

1. e-321=U+1F601 ANGUISHED FACE → WEARY FACE
Rationale: The KDDI source is an onomatopoeia which suggests the person is tired and whining.
2. e-326=U+1F606 EXPRESSIONLESS FACE → UNAMUSED FACE
Rationale: This follows a significant glyph change agreed to in Tokyo.
3. e-327=U+1F607 FACE WITH HEART-SHAPED EYES → HAPPY FACE WITH HEART-SHAPED EYES
4. e-329=U+1F609 WINKING FACE WITH STUCK-OUT TONGUE → KIDDING AND WINKING FACE WITH STUCK-OUT TONGUE
5. e-32A=U+1F60A FACE WITH STUCK-OUT TONGUE → KIDDING FACE WITH STUCK-OUT TONGUE
6. e-32C=U+1F60C FACE THROWING A KISS → HAPPY FACE THROWING A KISS
7. e-32D=U+1F60D FACE KISSING → HAPPY FACE KISSING
8. e-335=U+1F615 HAPPY FACE WITH WIDE MOUTH AND RAISED EYEBROWS → HAPPY FACE WITH WIDE MOUTH AND SMILING EYES
Rationale: "raised eyebrows" refer to older glyph designs than what was agreed to in FPDAM8
9. e-338=U+1F616 HAPPY FACE WITH OPEN MOUTH AND RAISED EYEBROWS → HAPPY FACE WITH OPEN MOUTH AND SMILING EYES
Rationale: "raised eyebrows" refers to older glyph designs than what was agreed for FPDAM8.
10. e-348=U+1F62B CAT FACE WITH OPEN MOUTH → HAPPY CAT FACE WITH OPEN MOUTH
11. e-34B=U+1F62E CAT FACE KISSING → HAPPY CAT FACE KISSING
12. e-34C=U+1F62F CAT FACE WITH HEART-SHAPED EYES → HAPPY CAT FACE WITH HEART-SHAPED EYES
13. e-34F=U+1F632 CAT FACE WITH TIGHTLY-CLOSED LIPS → SMART CAT FACE WITH TIGHTLY-CLOSED LIPS
14. e-357=U+1F63A PERSON RAISING ONE HAND → HAPPY PERSON RAISING ONE HAND
15. e-35B=U+1F63E PERSON WITH FOLDED HANDS → PLEADING PERSON WITH FOLDED HANDS

c. The U.S. does not recommend the following two name changes in N3711:

1. e-32E=U+1F60E FACE WITH MASK → SAD FACE WITH MEDICAL MASK
2. e-325=U+1F605 EXASPERATED FACE → IMPATIENT FACE

The U.S. recommends instead the following names:

1. e-32E=U+1F60E FACE WITH MEDICAL MASK
Rationale: While including “medical” in the name is good, "sad" seems too strong.
2. e-325=U+1F605 FACE WITH COLD SWEAT
Rationale: The KDDI meaning is “be hasty” and SoftBank literally means someone is off-balance or upset. Both reflect that someone is not calm and has lost his composure.

d. The U.S. requests a change to the glyph of e-33F=U+1F61D CONFOUNDED FACE.

The glyph has an irregularly-shaped squiggle over the head of this "confounded face". It should be removed.

e. The U.S. also requests a glyph change to e-34B=U+1F62E CAT FACE KISSING.

The glyph shows two mouth shapes: A heart representing "kissing", and also a simple mouth shape. The mouth shape looks like an error and should be corrected.

T.5. Emoji mapping data

a. The U.S. noted an error in N3728R, the mapping data for emoji. The character source emoji e-4EA REGIONAL INDICATOR SYMBOL LETTERS GB is currently mapped to U+1F1FA and U+1F1F0, which spell “UK.” However, the ISO 3166 code is “GB,” so the mapping data needs to be changed from:

1F1FA 1F1F0;;F3D1;FBB0

to:

1F1EC 1F1E7;;F3D1;FBB0

b. We propose adding the following text to the header of the EmojiSrc.txt file, N3728R:

This file provides mappings between UCS code points and sequences on one hand and Shift-JIS codes for cell phone carrier symbols on the other hand. Each mapping is symmetric ("round trip"), for equivalent UCS and carrier symbols or sequences. This file does not include best-fit ("fallback") mappings to similar but not equivalent symbols in either mapping direction.

Note: It is possible that future versions of this file will include additional data columns providing mappings for additional vendors.

T.6. CJK

a. The U.S. notes that 3 CJK sources are being mixed together in G_GFZB. G_GFZB is itself a union of ZhongHua ZiHai, Xiandai Hanyiu Cidian, and Ci-Hai. Xiandai Hanyiu Cidian and Ci-Hai already have their own source identifications, G_XC and G_CY, respectively. Instead of citing “G_GFZB,” the other two sources, G_XC and G_CY, should be cited.

b. The U.S. further strongly requests that all sources should have a numeric value, and this information be provided as soon as possible to the Project Editor. Although many G sources still allow a non-numeric value for the source (e.g., GBK, GCH, GFZ, GHZ), this information is nearly useless. Some sources with numeric values are easily available, such as Hanyu Dazidian.

Editorial Comments:

E.1. Malayalam

The U.S. requests an annotation be added to 0D4E MALAYALAM LETTER DOT REPH: “not used in reformed modern Malayalam orthography.”

Rationale: This annotation will make it clear that the letter is not used in the reformed orthographical context.

E.2. Emoticons

The U.S. recommends adding an annotation for e-814=U+1F3B6 MULTIPLE MUSICAL NOTES “x beamed eighth notes 266B”.