

INCITS/L2/10- 323

Date: August 13, 2010

Title: Comments accompanying the U.S. positive vote on ISO/IEC CD 10646 3rd edition (SC2 N4146)

Source: INCITS/L2

Action: Forward to INCITS

The U.S. National Body is voting Yes with comments on the CD for the third edition of ISO/IEC 10646.

Technical Comments

T.1. Sinhala Named Sequences

The U.S. recommends the following three named sequences for Sinhala be added:

SINHALA CONSONANT SIGN YANSAYA; 0DCA 200D 0DBA

SINHALA CONSONANT SIGN RAKAARAANSAYA; 0DCA 200D 0DBB

SINHALA CONSONANT SIGN REPAYA; 0DBB 0DCA 200D

T.2. Miao

a. The U.S. asks that an additional column (16F90-16F9F) be added to the Miao block so that it extends from U+16F00 to U+16F9F. The additional column is requested in order to accommodate character requests which are anticipated in the future.

b. The U.S. requests the Miao characters be re-ordered as documented in N3877 (L2/10-302), but with the name changes as noted below.

c. The U.S. recommends the names be changed from "MIAO LETTER" TO "MIAO VOWEL SIGN" for the characters in the range U+16F54 to U+16F7E.

Rationale: The name changes requested in (c), (d), and (e) will make the names more consistent with names elsewhere in the UCS.

d. The U.S. requests the following name changes:
from:

16F51 MIAO LETTER ASPIRATION

16F52 MIAO LETTER REFORMED VOICING

16F53 MIAO LETTER REFORMED ASPIRATION

to :

16F51 MIAO SIGN ASPIRATION

16F52 MIAO SIGN REFORMED VOICING

16F53 MIAO SIGN REFORMED ASPIRATION

e. The U.S. requests the following name changes:
from:

16F7D MIAO LETTER TONE RIGHT
16F7E MIAO LETTER TONE TOP RIGHT
16F7F MIAO LETTER TONE ABOVE
16F80 MIAO LETTER TONE BELOW

to (with new code points):

16F8F MIAO TONE RIGHT
16F90 MIAO TONE TOP RIGHT
16F91 MIAO TONE ABOVE
16F92 MIAO TONE BELOW

f. The U.S. requests the following 5 new characters be added, with glyphs and properties as documented in N3877 (L2/10-302), but with name changes as noted below for 16F56 and 16F5B:

16F0C MIAO LETTER YI TTA
16F12 MIAO LETTER YI NNA
16F31 MIAO LETTER YI DZHA
16F56 MIAO VOWEL SIGN AHH
16F5B MIAO VOWEL SIGN WO

T.3. Miscellaneous Symbols and Pictographs

The U.S. requests the names of the following 3 characters be changed, to use the spelling “POMMEE” instead of “POMMY”:

1F540 CIRCLED CROSS POMMEE
1F541 CROSS POMMEE WITH HALF-CIRCLE BELOW
1F542 CROSS POMMEE

Rationale: The spelling with -EE is more common than POMMY. The spelling POMMEE will also prevent possible confusion with a derogatory term POMMY. Supporting documentation is provided in document N3884 (L2/10-305).

T.4. Letterlike Symbols

The U.S. asks the following formal name alias be added to U+2118 SCRIPT CAPITAL P:
“WEIERSTRASS ELLIPTIC FUNCTION.”

Rationale: The current name is a misnomer, as it doesn’t describe the glyph (which is lowercase) or its function (Weierstrass elliptic function). A name alias will clarify the identity of this character.

T.5. Clause 25 Named UCS Sequence Identifiers

The US strongly prefers that the syntax for the data file containing the list of Named UCS Sequence Identifiers (nusi.txt) for 10646 maintain the same data format as the corresponding data file for the Unicode Character Database (NamedSequences.txt), or even that this data be maintained by normative reference to NamedSequences.txt, to prevent potential divergence between the two standards.

The notation for UCS Sequence Identifiers defined in Clause 6.6 is intended for use in descriptive text. In such descriptive text, an abbreviated format may be used, such as omitting the “U+” or the commas. Also, the use of angle brackets in such contexts is not problematical and helps in visual identification of the sequences. In data files, however, the notation may vary, and simpler formats that omit commas and angle brackets may be preferred. Omitting angle brackets simplifies data parsing and avoids conflicts with the use of angle brackets as part of the markup conventions for common markup languages such as HTML and XML. In binary-formatted data files the format for UCS Sequence Identifiers may be completely different.

Accordingly, the US NB requests that the documentation of the second field of nusi.txt be restored to the following:

- 2nd UCS sequence: (<space>)* (hhhh|hnhhh|hnhhhh) (<space> (hhhh|hnhhh|hnhhhh))+

If necessary, a note can be added explaining why this format is used in the data file, and how it can be systematically related to the syntax format defined for UCS Sequence Identifiers in Clause 6.6.