

Date: September 10, 2001

Title:	Comments accompanying the US positive vote on the FPDAM 1 to ISO/IEC 10646-1:2001
Source:	NCITS/L2
Action:	Forward to NCITS

The US is voting YES on the FPDAM ISO/IEC 10646-1:2000 with the following comments:

Technical Comments:

T.1 Item 1. Mathematical and other characters

Supplemental Arrows-A

The character name for U+27FF: LONG RIGHTWARDS ZIG-ZAG ARROW should be changed to LONG RIGHTWARDS SQUIGGLE ARROW to be consistent with the naming used for the characters U+21DC and U+21DD which are similar.

At minimum, the hyphen in ZIG-ZAG should be removed.

T.2 Item 1. Mathematical and other characters

Miscellaneous Mathematical Symbols-B

The glyph shapes of the characters U+29D8-29DB seem to contradict the left/right convention for glyph orientation. For example, is U+29D8 an opening notation (left) or closing as its shape may hint?

T.3 Item 1. Mathematical and other characters

Miscellaneous Mathematical Symbols-B

The characters U+29D1, U+29D2, U+29D4 and U+29D5 should be renamed as follows for consistency with other characters:

29D1 BOWTIE WITH LEFT HALF BLACK
 29D2 BOWTIE WITH RIGHT HALF BLACK
 29D4 TIMES WITH LEFT HALF BLACK
 29D5 TIMES WITH RIGHT HALF BLACK

T.4 Item 1. Mathematical and other characters

Miscellaneous Mathematical Symbols-B

The characters U+29FC-29FD should be renamed as follows for consistency with other characters (29E8-29E9):

29FC LEFT-POINTING CURVED ANGLE BRACKET
 29FD RIGHT-POINTING CURVED ANGLE BRACKET
 (Note the added hyphen in the names)

T.5 Item 1. Mathematical and other characters

Supplemental Mathematical Operators

The characters U+2AF7-2AF8 should be renamed as follows for consistency with other characters (2AA1-2AA2):

2AF7 TRIPLE NESTED LESS-THAN
 2AF8 TRIPLE NESTED GREATER-THAN

T.6 Item 1. Mathematical and other characters

Disunified Math Symbols

WG2 in its resolution M40.7 provisionally accepted 6 new math symbols and 2 new CJK symbols with instruction to the member bodies to review and comment. After further review, the US member body is in favor of adding these characters with the following names and code positions:

27E6 MATHEMATICAL LEFT WHITE SQUARE BRACKET
27E7 MATHEMATICAL RIGHT WHITE SQUARE BRACKET
27E8 MATHEMATICAL LEFT ANGLE BRACKET
27E9 MATHEMATICAL RIGHT ANGLE BRACKET
27EA MATHEMATICAL LEFT DOUBLE ANGLE BRACKET
27EB MATHEMATICAL RIGHT DOUBLE ANGLE BRACKET

FF5F FULLWIDTH LEFT WHITE PARENTHESIS
FF60 FULLWIDTH RIGHT WHITE PARENTHESIS
(The new allocation avoids the creation of new blocks)

T.7 Item 8. Collections for MES

MES-3A and MES-3B

The US National body is in favor of removing these two blocks because they are controversial. The current definition in the FPDAM document does not match the original repertoire from which these collections are originated from. This request doesn't preclude their further inclusion once the issue is settled.

T.8 Item 10. Annex B

B.1 List of combining characters/ Variation selectors

Remove the following entry:

0B83 TAMIL SIGN VISARGA

(This has been recognized as an error by Tamil language experts). The glyph in the chart pages should also have its dotted circle removed.

Add the following entries:

180B MONGOLIAN FREE VARIATION SELECTOR ONE
180C MONGOLIAN FREE VARIATION SELECTOR TWO
180D MONGOLIAN FREE VARIATION SELECTOR THREE
FE00 VARIATION SELECTOR-1

The above characters are combining characters per definition of sub-clause 4.12.

T.9 Item 10. Annex B

B.1 List of combining characters/ Variation selectors

The collection 103 VARIATION SELECTORS FE00-FE0F should be filled from FE01 to FE0F as follows:

FE01 VARIATION SELECTOR-2

...

FE0F VARIATION SELECTOR-16

And these characters should be added to the Annex B.1 as well. Finally, the collection should be shown with a '*' (full) in Annex A.

Finally, there is a need to create a new clause or sub-clause to describe the variation selectors. Currently the standard only contains an informative sub-clause about the Mongolian shaping selectors in F.2.5. It is necessary to create a normative description of the variation selectors as new ones are being introduced.

Variation selectors are both 'Special characters' (per clause 20 and 'Combining characters' (per clause 24). A sub-clause 24.5 could be added to the clause 24 Combining Characters to read as follows:

24.5 Variation selectors

Variation selectors are combining characters following immediately a specific base character to indicate a specific variant form of graphic symbol for that character. Some variation selectors are specific to a script, like the Mongolian free variation selectors, others are used with various other base characters like CJK characters or

Mathematical symbols. Variations selectors following other characters have no effect on the selection of the graphic symbol for that character. The base characters defined for use with the variant selector are given in the following table:

<u>Sequence (UID notation)</u>	<u>Description of variant appearance</u>
<2268, FE00>	less-than and not double equal - with vertical stroke
<2269, FE00>	greater-than and not double equal - with vertical stroke
<22DA, FE00>	less-than above slanted equal above greater-than
<22DB, FE00>	greater-than above slanted equal above less-than
<2272, FE00>	less-than or similar – following the slant of the lower leg
<2273, FE00>	greater-than or similar - following the slant of the lower leg
<2A9D, FE00>	similar - following the slant of the upper leg - or less-than
<2A9E, FE00>	similar - following the slant of the upper leg - or greater-than
<2AAC, FE00>	smaller than or slanted equal
<2AAD, FE00>	larger than or slanted equal
<228A, FE00>	subset not equals - variant with stroke through bottom members
<228B, FE00>	superset not equals – variant with stroke through bottom members
<2ACB, FE00>	subset not two-line equals - variant with stroke through bottom members
<2ACC, FE00>	superset not two-line equals - variant with stroke through bottom members
<2A3B, FE00>	interior product - tall variant with narrow foot
<2A3C, FE00>	righthand interior product – tall variant with narrow foot
<2295, FE00>	circled plus with white rim
<2297, FE00>	circled times with white rim
<229C, FE00>	equal sign inside and touching a circle
<222A, FE00>	union with serifs
<2229, FE00>	intersection with serifs
<2293, FE00>	square intersection with serifs
<2294, FE00>	square union with serifs

(Information about Mongolian and CJK should be added to the table as well)

NOTE 1: Clause F.2.5 contains additional information about the Mongolian free variation selectors.

NOTE 2: The variation selector only selects a different *appearance* of an already encoded character. It is not intended as a general code extension mechanism. Only the sequences specifically defined in this annex are sanctioned for standard use, all other sequences are undefined. No sequences containing combining characters or composite characters will be defined.

(end of addition)

T.10 (not related to a FPDAM item)

A normative reference should be added in ISO/IEC 10646-1 pointing to the Unicode Bidirectional Behavior (clause 3.12 of the Unicode Standard 3.0). It is impossible to implement bidirectional text processing without a detailed description of the bidirectional algorithm.

T.11 Item 14. Compatibility Ideographs and source separation rules

22.2 Source references for CJK Compatibility Ideographs

The new compatibility character 0FA45 shows a mapping as follows:

0FA45 06982 J3-7624

(Indicating a mapping to the Unified character 06982 with source reference JIS X 213:2000 level-3)

The US national body is in favor of mapping it instead to the Unified character 069EA. The following two lines show the C(G-T), J and K graphic symbols shown in typical fonts for these two characters:

06982 概 (SimSun) 概 (MingLiu) 概 (MS Mincho) N/A (Batang)

069EA 概 (SimSun) 概 (MingLiu) N/A (MS Mincho) 概 (Batang)

The closest character in shape to the original JIS X 213 characters is located in the second line. It should be noted that the Taiwanese source (MingLiu) shown here differs significantly from the graphic symbols shown in ISO

10646-1:2000 for the similar source reference. At minimum the status of these two characters is unclear, and before mapping a compatibility character into one of them, the original encoding rationale of these two characters should be better understood.

The two following pages show the two pages from IS/IEC 10646-1:20000 CJK character in context.

---end of US comment---
(Annex follows)