ISO/IEC JTC 1/SC 2

CODED CHARACTER SETS

SECRETARIAT: JAPAN (JISC)

DOC TYPE: Proposal for Project Subdivision

TITLE: Subdivision Proposal on JTC 1.02.18.01 for Amendment 1: Encoding space limitations and repertoire additions to ISO/IEC 10646-1: 2000, Universal Multiple-Octet Coded Character Set (UCS) -- Part 1: Architecture and Basic Multilingual Plane

SOURCE: ISO/IEC JTC 1/SC 2/WG 2

PROJECT: JTC 1.02.18.01

STATUS: Based on RESOLUTION M38.14 (Future amendment to 10646-1: 2000) WG2 requests SC2 approval for sub-division of work to prepare a future amendment to ISO/IEC 10646-1:2000. SC 2 members are requested to vote on this proposal as soon as possible but not later than 2000-08-30.

ACTION ID: LB

DUE DATE: 2000-08-30

DISTRIBUTION: P, O and L Members of ISO/IEC JTC 1/SC 2
WG Conveners and Secretariats
Secretariat, ISO/IEC JTC 1
ISO/IEC ITTF

NO. OF PAGES: 3

ACCESS LEVEL: Open

WEB ISSUE #:   088

Proposed Working Draft

Subject: Subdivision of work – ISO/IEC 10646-1:2000

Status: Based on RESOLUTION M38.14 (Future amendment to 10646-1: 2000) WG2 requests SC2 approval for sub-division of work to prepare a future amendment to ISO/IEC 10646-1:2000.

Proposed Project Number: JTC 1.02.18.01.01.00.02

Title: ISO/IEC 10646-1, 2000: Amendment 1: Encoding space limitations and repertoire additions.

Scope:

1. **Remove the provision for private use groups and planes beyond plane 16 – Resolution M38.6.**

2. **Repertoire additions to BMP include:**

   a. **Approximately 951 Math symbols – Resolution M38.9**

   b. **Additional 5 Yi Radicals – Resolution M38.8**

   c. **Additional symbols from JIS X0213 – Resolution M38.10**

   d. **Additional 61 (approx) CJK compatibility ideographs from JIS X0213 – Resolution M38.11**

   e. **Additional Collection Identifiers for European sub repertoires – Resolution M38.13**

   f. **Additional Collection Identifiers for identifying collections that include base characters and combining sequences.**

   g. **Additional collection identifiers identified in draft of ISO/IEC 10646-2 – Resolution M38.4, the disposition of comments on the CD ballot for ISO/IEC 10646-2.**

   h. **Additional modifications based expected disposition of comments on ballot results of FCD ISO/IEC 10646-2.**

   i. **Miscellaneous repertoire additions such as Cyrillic Sami, Khmer, Peso Sign, Square Root, Property Line and German Penny Symbol - Resolution M38.7**

Rationale:

1. **Removing the provision for private use groups and planes beyond plane 16 makes UTF-8 and UTF-16 interoperable.**

2. **Repertoire additions are based on requests from national body and other experts in WG2 to meet their user community requirements.**

Target dates: WD – 2000-10; FPDAM –2001-02; FDAM 2001-08; AMD 2002-03

Relevant Resolutions text from N2204R – Meeting 38 in Beijing

RESOLUTION M38.4 (CD 10646-2):

Unanimous

WG 2 accepts the disposition of comments for the ballot on CD 10646-2 in document N2217, which reflects the results of discussion at this meeting M28, and instructs its project editor to prepare the final text of FCD 10646-2, including additional compatibility ideographs accepted for inclusion in Plane 2 at this meeting (per resolution M38.2 above), with assistance from the contributing editors, and to forward these documents to SC 2 secretariat for further processing in May 2000, with unchanged target dates.

RESOLUTION M38.6 (Restriction of encoding space):

Unanimous

WG2 accepts the proposal in document N2175 towards removing the provision for Private Use Groups and Planes beyond Plane 16 in ISO/IEC 10646, to ensure internal consistency in the standard between UCS-4, UTF-8 and UTF-16 encoding formats, and instructs its project editor prepare suitable text for processing as a future Technical Corrigendum or an Amendment to 10646-1: 2000.

RESOLUTION M38.7 (Additions to the BMP - Miscellaneous):

Unanimous

WG2 accepts to encode the following in the BMP of ISO/IEC 10646:

1) 17DD - KHMER SIGN LAAK, with the glyph as shown in document N2164.

2) 8 Cyrillic Sɹmi characters, with the shapes as shown on page 4 of document N2173:

     048A - CYRILLIC CAPITAL LETTER SHORT I WITH TAIL
     048B - CYRILLIC SMALL LETTER SHORT I WITH TAIL
     04C5 - CYRILLIC CAPITAL LETTER EL WITH TAIL
     04C6 - CYRILLIC SMALL LETTER EL WITH TAIL
     04C9 - CYRILLIC CAPITAL LETTER EN WITH TAIL
     04CA - CYRILLIC SMALL LETTER EN WITH TAIL
     04CD - CYRILLIC CAPITAL LETTER EM WITH TAIL, and,
     04CE - CYRILLIC SMALL LETTER EM WITH TAIL.

3) 20B1 - PESO SIGN, with the shape as shown on page 1 of document N2161 (a P with two horizontal bars)

4) 237D - SQUARE FOOT, and,
     234A - PROPERTY LINE
     with shapes similar to what appear in document N2184.

*NOTE: the above should be corrected to **23CD** and **214A** respectively. The error was discovered after the meeting by Dr. Ken Whistler -- Umamaheswaran.*

5) 20B0 - GERMAN PENNY SYMBOL with shape as shown in document N2188.

WG2 further instructs its project editor to add these characters to the list of characters accepted for processing beyond the 2nd edition of ISO/IEC 10646-1 (based on resolution M36.19 - New character bucket M36).

RESOLUTION M38.8 (Yi radicals):

Unanimous

WG2 provisionally accepts the five Yi radicals, their code positions, names and shapes, as shown in document N2207, subject to verification and feedback on these characters from the Chinese Yi experts.

RESOLUTION M38.9 (Math Symbols):

Unanimous

WG2 provisionally accepts the 951 Mathematical Symbols for encoding in the BMP, as proposed in document N2191R. The US national body is invited to prepare a revised proposal reflecting feedback received from the experts group on the subject before the next meeting M39 in September 2000. WG2 intends to include this proposal in the working draft of the next Amendment to 10646-1: 2000 at the next meeting.

RESOLUTION M38.10 (Symbols from JIS X 0213):

China Abstains

WG2 provisionally accepts the proposed symbols described under items 1 through 12 in document N2195 for future encoding in the BMP. Their names, shapes and code positions are to be synchronized with the processing of Mathematical Symbols proposal (see resolution M38.9 above), based on feedback to the experts' group on this topic.

RESOLUTION M38.11 (CJK compatibility ideographs from JIS X 0213):

China Negative

WG2 accepts the proposal for 61 compatibility ideographs from Japan for future encoding in the BMP, at positions FA30 to FA6B, with the shapes and names as shown in document N2197. Japan is to provide additional information on these compatibility ideographs suitable for inclusion in the standard per discussion at meeting 38 (see also document N2142). WG2 further instructs its project editor to add these characters to the list of characters accepted for processing beyond the 2nd edition of ISO/IEC 10646-1 (based on resolution M36.19 - New character bucket M36).

RESOLUTION M38.13 (Collection identifiers for Multilingual European Subsets (MES)):

Unanimous

WG2 accepts the request and instructs its editor to assign four collection identifiers for each of the Multilingual European Subsets - MES-1 (fixed collection), MES-2-(fixed collection), MES-3A (non-fixed collection) and MES-3B (fixed collection), as requested in document N2211, for inclusion in a future amendment to ISO/IEC 10646-1: 2000.