ISO/IEC JTC1/SC2/WG2
Coded Character Set
Secretariat: Japan (JISC)

Doc. Type: **Disposition of comments**

Title: **Proposed Disposition of comments on SC2 N 3585 (PDAM text for Amendment 1 to**
ISO/IEC 10646-2:2001)

Source: **Michel Suignard (project editor)**
Project: JTC1 02.18.02
Status: For review by WG2
Date: 2001-05-21
Distribution: WG2
Reference: SC2 N3597, N3603, N3605
Medium: Paper

Comments were received from Canada, Germany, Ireland, DPRK, ROK, Sweden and USA. The following document is the draft disposition of those comments. The disposition is organized per country.

> Note - The full content of the ballot comments (minus some character glyphs) have been included in this document to facilitate the reading. The dispositions are inserted in between these comments and are marked in **<u>Underlined Bold Serif text</u>**, *with explanatory text in italicized italic serif*.

<u>All the received comments were accompanying Yes votes.</u>

Canada: Yes with comments:

Technical comment:

Comment i – Variation selector supplement collection number.
On page 1, collection number 3002 for VARIATION SELECTORS SUPPLEMENT has to be changed to 3003 (or something else), because the number 3002 is already assigned to ALTERNATE FORMAT CHARACTERS in 10646-2: 2001 (page 382, Annex A.1)

Accepted
Collection changed to 3003.

Editorial Comments:

Comment ii:
Page 2 - text proposed for the 6 first lines within the linked file appears under the wrong column, and at first look is confusing.
Accepted

Comment iii:
Pages 64, 65 - Glyphs for Variation Selectors (also for the Variation Selectors in Part 1) - Having VSX only within the dashed square does not permit glyphic distinction between the various VS-s. Suggest moving the VS to the top half and inserting the numbers like 017, 018, .. 255 in the lower half of each box.
Accepted in principle

Comment iv:
Pages 64, 65 - Tables 202 and 203: Request moving column E018 from Table 202 to 203, to make the second table begin on a half-row boundary. The corresponding names lists should also be adjusted accordingly
Accepted in principle
The block is moved to E0100-E01EF with characters moved accordingly.

Germany: Yes with comments:
Technical comments:

Comment 1: Page v, Foreword:
Check the change against ISO's stated intention to improve the involvement of industrially less advanced countries in standardization. Experts from these countries may not have the means to access machine readable information. In this case, it might be necessary for ISO Central Secretariat to offer a version with all parts printed out.
Follow any advice from ISO Central Secretariat on this issue.
Partially accepted
The Foreword can identify explicitly by using filenames which parts of the standard are intently mostly as machine readable files. Because these materials are still printable (although not page formatted to preserve their machine readable feature) this does not constitute a barrier to access to industrially less advanced countries.
Recently, this standard has been distributed in a CD-ROM media form which contains all files which can be printed as desired by the recipients. CD-ROMs are much cheaper to reproduce than large documents and can be considered improvements to access to the standard.
The foreword notice will be changed as follows:

The standard contains material which is only available to users who obtain their copy in a machine-readable format. That material consists of the following printable files:
CJKB_SR.txt
CJKC_SR.txt
SMPnames.txt
SSPnames.txt

[…]

Comment 2: Linear B and Aegean Numbers:
Germany may comment on these scripts and characters in a later stage of the ballot process.
<u>Noted</u>

Comment 3: Osmaniya:
Sources such as <u>http://www.proel.org/alfabetos/somali.html</u> and
<u>http://www.omniglot.com/writing/somali.htm</u> gives significantly different versions of the
Osmaniya alphabet (primarily concerning glyph and order of presentation). Germany does not have the
expertise to evaluate the different presentations on the Osmaniya alphabet and therefore simply requests that
the editor check the correctness of the listed characters against the cited sources.
<u>Accepted in principle</u>
The document WG2 N2361R is a much more authoritative source for these characters and is based on a
Somali publication by the great-son of the original author.

<u>Ireland: Yes with comment:</u>

Technical comments

T1. Tables 202 and 203.
Change "VARIATION SELECTORS" to "VARIATION SELECTORS SUPPLEMENT" in the block name as
per the item at **Page 382, Annex A.1** on page 1. (The block name "VARIATION SELECTORS" is already in
use for the block at U+FE00.)
<u>Accepted</u>

Editorial comments

E1. Page 382, Annex A.1.
Change "AEGAN" to "AEGEAN".
<u>Accepted</u>

E2. Page 382, Annex A.2.
Change "AEGAN" to "AEGEAN".
<u>Accepted</u>

E3. 2. New tables.
Change "Lineam B Ideograms" to "Linear B Ideograms".
<u>Accepted</u>

E4. 2. New tables.
Change Ugaratic" to "Ugaritic".
<u>Accepted</u>

E5. Table 172 - Row 104: Osmanya.
The font is rather rough and should be improved.
<u>Accepted in principle</u>
The editor needs a new font to accommodate this

E6. Table 202 - Row E01: Variation Selectors Supplement.
The identifying number should be added to all of the glyphs here.
<u>Accepted in principle</u>
The editor needs a new font to accommodate this

<u>Korea, Democratic People's Republic (DPRK), Yes with comment</u>

Technical Comments

T1. CJK Mapping table.
In the file 'CJKB_SR.txt', map the code of <KP1-6134> to <U2564A>.
Accepted
Note that this is a new entry; KP1-6134 didn't exist in the current CJKB_SR.txt file. Reference document WG2 N2427.

T2. CJK Compatibility Mapping table.
In C.2 of page 2 in ISO/IEC 10646-2:2001/PDAM 1, add following text
`'Hanja KP sources: 29-36 OCTET'`
after
`'...Hanzi H Sources: 22-28 OCTET'`
Accepted

T3. CJK Compatibility Mapping table.
In the file 'CJKB_SR.txt', add <U26C2A>
[Note]
Line of <U26C2A> is empty. Although there is no mapping in this line, we consider that the line should not be empty.
Accepted in principle
From looking at the file, it looks like the line U262CA is not missing, but the previous line is, it should read:
`26C29          J4-2544`
(Assuming it has no KP1 reference, and both lines have mappings)
The point has also been raised by ROK, comment E.2

T4. CJK Compatibility Ideographs extension
Add 16 Compatibility ideographs into the CJK Compatibility Ideographs Supplement.
[Note]
> M41.10 WG2 invites DPR of Korea to examine their proposal for 160 CJK compatibility characters towards reducing the total number to maximum of 151 with the goal of being able to fit them into the available space in the BMP. WG2 further invites DPR of Korea to input this revised proposal as part of their ballot comments to Amendment 2 of IO/IEC 10646-1:2000.

According to this resolution, we reduced our compatibility ideographs for BMP to 144 characters. We already sent the document on 16 compatibility ideographs for SMP to Mr. Suignard, Chief editor of WG2 on 6th Dec, 2001 by E-mail. We also sent the TrueType Font and code mapping table of the 16 Compatibility ideographs to Mr. Asumusu [sic], representative of Unicode Consortium.

Accepted in principle
As a result of disposing 10646-1:2000 Pdam2 ballot comments, 37 of the initial 160 characters are mapped to existing CJK compatibility characters in plane 2.

Editorial comments

E1. Replace KPS by KP in clause 7.2
In 7.2 of page 1 in the ISO/IEC 10646-2:2001/PDAM1, replace
'Insert the new Hanja KP<u>S</u> source references after the Hanja K sources as follows:
Hanja KPS sources are
KP0 KPS 9566-97
KP1 KPS 10721-2000'
By
'Insert the new Hanja KP source references after the Hanja K sources as follows:
Hanja KP source is
KP1 KPS 10721-2000'
Accepted

E.2 Compatibility CJK mapping table
In .1 of page 2 in the ISO/IEC 10646-2:2001/PDAM1, replace
'48-55 octet: Hanja KP sources (KP0-hhhh) or (KP1-hhhh)'
'48-55 octet: Hanja KP sources (KP1-hhhh)'
<u>Accepted</u>

Korea, Republic of (ROK), Yes with comment

Editorial Comments

E1. Source references for CJK unified ideographs.
P.384, clause C.1 Source references for CJK Unified Ideographs Extension B editorial - As noted in the editor's note, we request that "K-" in the source reference be changed to "K4-".
<u>Accepted in principle</u>
IRG has been requested to update its SuperCJK database to include the K4 reference notation and extract the relevant source reference information.

E2. Source references for CJK unified ideographs.
- In source reference file CJKB_SR.txt, line for Han ideograph U+26C29 seems missing. If this is an error, we need to add one line. If no character has been assigned to this code position, we suggest that we add a comment indicating that fact.
<u>Accepted</u>
Also found by DPRK, comment T3. The line is not empty.

Sweden: Yes with comments:

Technical comments:

SE 1.
page 1: "3.x" --> "4.0"?
<u>Accepted</u>

SE 2.
page 2: the two-column to one-column switch is not marked in any way, and it is confusing to read the first column, not knowing that there is a two- to one-column switch.
<u>Accepted</u>

SE 3. (Data file format)
page 2: A semicolon-separated format for the data file would be preferable, and would avoid the byte counting oddity. Say explicitly that the encoding of the data file is ISO/IEC 646 IRV with LINE FEED as end of line mark.
<u>Partially accepted</u>
A similar comment was made during the creation of Part which introduced the concept of linked data files. A similar comment was made for Part 2 when the files were introduced. These data files are mechanically generated by IRG and it would be very unwise to manually change the format of these as it could introduce errors. The format is easy to import by database tools as it stands.
The second part of the comment is already covered by the mention to ISO/IEC 646-IRV characters (section C1 in page 384). The only part that could be added concerns 'LINE FEED as end of mark'. Note however that all clauses that use the same languages in both parts need to be updated accordingly. (another argument for merging part 1 and 2)

SE 4.
U+100D5, U+100D6, and U+100FA appear to be for characters where only partial glyphs are known. It should be stated that these codes are for the (or one of the) character(s), whether the known glyph is partial or not. The partiality per se must be noted by other means (like markup).

Page 5 of 6

Noted
U+100BF belongs to the same category. The dots are considered part of the glyphs.

SE 5.
U+E0110, etc.: more than 200 variation selectors is excessive. There is no planned used for that many variation selectors. Only much smaller number should be allocated.
Withdrawn

USA: Yes with comments:

Editorial Comments:

E.1 Reference to the Unicode Standard
Replace the reference to Unicode 3.x by 'Unicode 4.x'
Accepted in principle
It should be 4.0 (not 4.x)

E.2 Aegean
Correct the misspelling in page 1 in the entries (at least two occurrences)
Accepted

E.3 New tables (page 3)
Correct the name of table 2 (Linear)
Correct the name of table 6 (Ugaritic)
Accepted

E.4 Page numbers and pages
As reported in the EDITOR'S NOTE at the end of page 3.
Accepted

E.5 Table 202 Row E01: Variation Selectors
All glyphs have the same 'VSX'; they should be shown as going from 'VX17 to 'VSX256'.
Accepted

[end]