The US National Body is voting Yes with comments on the following SC2 ballot.

SC2N3943: Information technology -- Universal Multiple-Octet Coded Character Set (UCS) -- AMENDMENT 4: Lanna, Cham, Game Tiles, and other characters

**Technical Comments:**

**T.1 Source reference for existing CJK characters (page 2)**

The WG2 Resolution M50.20 (dis-unification of CJK Unified ideograph 4039) was not completely conveyed in amendment 4. In addition to the removal of T4-3946, it also needs to be replaced by T6-4B7A. Furthermore, the linked file CJKU_SR.txt has none of these changes for the U+04039 entry. Note that the U+09FC3 is correct, and that this also needs to be fixed in the proposed amendment 5 which contains the CJKU_SR.txt file.

**T.2 New collection numbers (page 3)**

In page 3 of the amendment, the collections number for the proposed scripts are incorrect because they re-use number already allocated to scripts added by amendment 3. So the text after ‘112 ARABIC SUPPLEMENT’ should read:

In the list of collection numbers and names, after 139 REJANG

insert new entries as follows:

<table>
<thead>
<tr>
<th>Collection Name</th>
<th>Collection Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lanna</td>
<td>1A20-1AAF</td>
</tr>
<tr>
<td>Cyrillic Extended-A</td>
<td>2DE0-2DFF *</td>
</tr>
<tr>
<td>Cyrillic Extended-B</td>
<td>A640-A69F</td>
</tr>
<tr>
<td>Cham</td>
<td>AA00-AA5F</td>
</tr>
</tbody>
</table>

These collection numbers are again used in the next page (ref Note 3 about keywords), and therefore they should be updated there as well.

**T.3 Deprecation of existing collections (page 3)**

The collection 271 COMBINING CHARACTERS B-2 is specified by direct referencing annex B.2 that this amendment is removing. Therefore it needs to be deprecated with wording similar to other deprecated collections (such as collections 400 and 500). The text for the definition of collection 270 COMBINING CHARACTERS and 1900 SMP COMBINING CHARACTERS also needs to be changed to point to annex B instead of annex B.1.

**T.4 Shan characters (Myanmar)**

The US is in favor of the re-ordering of the Shan characters as shows in WG2 N3277 which are already in the amendment (these are shown in blue in the document) to make room for additional characters (shown in yellow) in the future. N3277 also proposes a glyph change for 107A (formerly 1078) and name change for 108C (formerly 1085) from ‘MYANMAR SIGN SHAN COUNCIL TONE-4’ to ‘MYANMAR SIGN SHAN COUNCIL TONE-3’ which are agreeable by the US. In addition, the glyph for 1040 MYANMAR DIGIT ZERO should reflect actual use.
If the glyph is identical to the one used for 101D MYANMAR LETTER WA, there should not be an artificial difference as currently shown in N3277.

**T.5 Malayalam Chilu characters**

The US NB does not support a name change for 0D7C MALAYALAM LETTER CHILLU RR.

**T.6 Latin Epigraphic characters**

The US NB does not support the addition of lower case forms to these characters.

**T.7 Mirrored characters**

After further study the US NB is in favor of undoing the change made by Amendment 3 concerning the mirrored property of the following characters.

- 2018 LEFT SINGLE QUOTATION MARK
- 2019 RIGHT SINGLE QUOTATION MARK
- 201A SINGLE LOW-9 QUOTATION MARK
- 201B SINGLE HIGH-REVERSED-9 QUOTATION MARK
- 201C LEFT DOUBLE QUOTATION MARK
- 201D RIGHT DOUBLE QUOTATION MARK
- 201E DOUBLE LOW-9 QUOTATION MARK
- 201F DOUBLE HIGH-REVERSED-9 QUOTATION MARK
- 301D REVERSED DOUBLE PRIME QUOTATION MARK
- 301E DOUBLE PRIME QUOTATION MARK
- 301F LOW DOUBLE PRIME QUOTATION MARK

As a result these characters should not be mirrored in bidirectional context.

**Editorial Comments:**

**E.1 Definition of UNICODE 5.1 in A.6.6 (page 4)**

Correct the definition as follows:

308 The fixed collection UNICODE 5.1 is arranged by planes as follows.

**E.2 Identification of UTF-16 (page 7)**

In the updated paragraph replace ‘sequence and note’ with ‘sequence’

**E.2 Identification of UTF-8 (page 7)**

In the updated paragraph replace ‘by following’ with ‘by the following’

**E.2 Identification ASN.1 character abstract syntaxes (page 7)**

Fix the Note 4 and following paragraph as follows:

NOTE 4 – As an example, the object identifier for the subset comprising the collections BASIC LATIN, LATIN-1 SUPPLEMENT, and MATHEMATICAL OPERATORS is:

{iso standard 10646 (0) level-3 (3) collections (1) 1 2 39}

ISO/IEC 8824 also specifies object descriptors corresponding to object identifier values. For an unrestricted repertoire, the corresponding object descriptor is as follows:

---End of US comments