ISO/IEC JTC 1/SC 2
Coded Character Sets
Secretariat: Japan (JISC)

DOC. TYPE	Summary of Voting/Table of Replies
TITLE	Summary of Voting on ISO/IEC JTC 1/SC 2 N 3982 : ISO/IEC 10646: 2003/FPDAM 5, Information technology -- Universal Multiple-Octet Coded Character Set (UCS) -- AMENDMENT 5: Tai Tham, Tai Viet, Avestan, Egyptian Hieroglyphs, CJK Unified Ideographs Extension C, and other characters
SOURCE	SC 2 Secretariat
PROJECT	JTC1.02.10646.00.05
STATUS	This document is forwarded to WG 2 for disposition of comments. WG 2 is requested to prepare a disposition of comments report, revised text and a recommendation for further processing.
ACTION ID	FYI
DUE DATE	
DISTRIBUTION	P, O and L Members of ISO/IEC JTC 1/SC 2 ; ISO/IEC JTC 1 Secretariat; ISO/IEC ITTF
ACCESS LEVEL	Def
ISSUE NO.	304
FILE	**NAME** 02n4004.pdf **SIZE (KB)** **PAGES** 16

<table>
<tr><td colspan="6" align="center">Summary of Voting on
ISO/IEC JTC 1/SC 2 N 3982 :
ISO/IEC 10646: 2003/FPDAM 5, Information technology -- Universal Multiple-Octet Coded Character Set (UCS) -- AMENDMENT 5: Tai Tham, Tai Viet, Avestan, Egyptian Hieroglyphs, CJK Unified Ideographs Extension C, and other characters</td></tr>
</table>

Q1 : FPDAM Consideration

	Q1			No response	Comments Reasons
	Approve	Disapprove	Abstention		
P-Member					
Belgium				X	
Canada	X				
China	X*				Attachment 1
Egypt				X	
Finland	X				
France				X	
Germany	X				
Greece				X	
Iceland				X	
India	X				
Indonesia	X				
Iran, Islamic Republic of				X	
Ireland		X#			Attachment 2
Italy	X				
Japan		X#			Attachment 3
Kazakhstan				X	
Korea, Democratic People's Republic				X	
Korea, Republic of		X#			Attachment 4
Mongolia				X	
Norway	X				
Poland	X				
Romania	X				
Russian Federation				X	
Serbia				X	
Sweden			X		
Thailand			X		
Tunisia				X	
United Kingdom		X#			Attachment 5
USA		X#			Attachment 6
Total (29)	**10**	**5**	**2**	**12**	

*: Approve with comments
#: Acceptance of the reasons and appropriate changes in the text will change the vote to approval.

O-Member					
Sri Lanka	N/A	N/A	N/A		Attachment 7

China votes YES to SC2n3982 with the following comments on script TAI THAM:

1. Names of digits
   China requires to change character names of U+1A80 through U+1A89. The names should be changed to TAI THAM HORA DIGIT ZERO … NINE, as agreed on ad hoc in Chiang Mai early this year. (WG2n3389 Item 2)

2. Fonts
   China requires to change fonts to Khuen/Lue style, as agreed on ad hoc in Chiang Mai early this year. (WG2n3389 Item 4)

3. Moving some characters
   China requires to move U+1A56 to U+1A50, move U+1A50 through U+1A55 to U+1A51 through U+1A56, as agreed on ad hoc in Chiang Mai early this year. (WG2n3389 Item 6)

4. Representative Glyph for U+1A31
   China requires to change the Representative Glyph for U+1A31 to Khuen/Lue style, as agreed on ad hoc in Chiang Mai early this year. (WG2n3389 Item 7)

5. Remove Language Specific Letters
   China requires to remove some language specific letters, as agreed on ad hoc in Chiang Mai early this year. (WG2n3389 Item 8)

For details, read WG2n3389 please.

Irish comments on FPDAM-5 for ISO/IEC 10646:2003
Reference: SC2 N3982
Closes: 2008-03-20
Date: 2008-03-14

Ireland **disapproves** the draft with the technical and editorial comments given below. Acceptance of these comments and appropriate changes to the text will change our vote to approval.

Technical comments

T1. **Page 9, Row 0D8, Sinhala.** Following on from discussions with members of the Sri Lankan standards body regarding a need for further consultation and study of the Sinhala digits, Ireland requests that the characters under ballot in columns 0D8E and 0D8F be removed from FPDAM 5 and placed on a subsequent ballot.

T2. **Page 14, Row 110: Hangul Jamo.** Ireland requests that the character name at U+11FD HANGUL JONGSEONG KIYEOK-KHIEUK be corrected to HANGUL JONGSEONG KIYEOK-KHIEUKH since the final -H is missing.

T3. **Page 16, Row 1A2: Tai Tham.** With reference to ISO/IEC JTC1/SC2/WG2 N3379 "Tai Tham Ad-hoc Meeting Report", Ireland requests that the character set and code table in FPDAM 5 be replaced by that in N3379.

T4. **Page 16, Row 1A2: Tai Tham.** With reference to ISO/IEC JTC1/SC2/WG2 N3384 "Tai Tham Subjoined Variants", Ireland requests that U+1A5D TAI THAM CONSONANT SIGN BA and U+1A5E TAI THAM CONSONANT SIGN SA be added to FPDAM 5.

T5. **Page 14, Row 110: Hangul Jamo Extended-A.** Ireland requests that the character names at U+A96E HANGUL CHOSEONG RIEUL-KHIEUK and U+A973 HANGUL CHOSEONG PIEUP-KHIEUK be corrected to HANGUL CHOSEONG RIEUL-KHIEUKH and HANGUL CHOSEONG PIEUP-KHIEUKH respectively since the final -H is missing.

T6. **Page 40, Row 10B0: Avestan.** Ireland reiterates its support for the retention of U+10B38 AVESTAN SEPARATION POINT, a character distinct from FULL STOP which is also used in Geldner's edition of the Avesta. We have reviewed N3336 but do not agree that workarounds like lowering U+2E37 are the correct way to encode the character in question. The complaint about the baseline drawing in N3197 is not well-made; those lines were indicative and were constructed by shifting out-of-text graphic elements (two horizontal hairlines) around in Quark XPress. The AVESTAN SEPARATION POINT does sit on the baseline; there is no real "discrepancy between the two documents regarding the location" of the character. We understand that the U.S. National Body's position is that there are too many dots in the standard; we maintain our view that script-specific punctuation is preferable, in general, to over-unification. Ireland stands by its comments to PDAM 5 on this point, which is in accord with the comments of Professor Jost Gippert, one of the specialists who worked on the encoding of Avestan. We would not agree to remove this character from FPDAM 5.

Editorial comments

E1. **Page 12, Row 110: Hangul Jamo.** Ireland notes discrepancies between the font used in Unicode 5.0 and the font used in FPDAM 5. For instance, U+110C HANGUL CHOSEONG CIEUC and U+119E HANGUL JUNGSEONG ARAEA differ between them. Ireland requests that WG2 work with the Korean delegates, the Unicode Liaison, and other interested National Bodies in determining the preferred shapes for this and similar characters.

E2. **Page 24, Row A96: Hangul Jamo Extended-A.** The same comment applies here as in E1.

E3. **Page 27, Row D7B: Hangul Jamo Extended-B.** The same comment applies here as in E1.

E4. **Page 48, Row 1300: Egyptian Hieroglyphs.** Ireland notes that the font is missing in the annotation for U+1309D EGYPTIAN HIEROGLYPH D036, and U+1313F EGYPTIAN HIEROGLYPH G001. There is also an extra comma in the note at U+133ED EGYPTIAN HIEROGLYPH Z004.

Japan disapproves SC2 N3982 (ISO/IE10646 FPDAM 5) with the following comments. Japan will change its vote if those comments are addressed appropriately.

Technical comments:

JPT1: Japan supports the consensus on the IRG #29 meeting regarding so-called ARIB ideographs.   Take a character, J_ARIB-757D, currently on FA6D, out of the compatibility ideographs and move it into the unified ideographs.   See IRG document N 1347R for details. The amendment text and related files should be updated appropriately.   (Also see JPE1.)

JPT2: The draft amendment text contains several references to the Unicode Standard 5.2, including the collection 309.   Japan is not sure it will be published on time.   Japan wants to know the plans and schedules of Unicode 5.2 publishing from the Unicode Consortium. Japan wants to remove those references from the amendment texts unless a satisfactory plan is provided.

JPT3: On the right column of page 3, there is a new text for 34.2.   The third item reads "Normative character alias (one preceded by "   ") which is a formal, unique, and stable alternate name for a character."   The current sentence seems redundant for a normative clause.   Considering these items are lead by a sentence "The following information items are normative" and a NOTE follows it, just a simple list item text such as "Character alias preceded by   " is enough.

Editorial comments:

JPE1: On the left column of page 5, regarding the part of the code chart for new CJK UNIFIED IDEOGRAPHS (9FC4 and 9FC5), the current chat uses low quality images of the new characters.   Use TrueType font supplied by Japan to typeset them.   (Also see JPT1.)

JPE2: On the left column of page 3, there is a new text for 27.3, "However, CJK Unified Ideographs Extension C uses a different format..."   The world "However" seems inappropriate here and should be removed.

Other comment:

JPO1: On the right column of page 2 to the next page, there is an "Editor's note".   It is unsure what it means.   Japan wants to see clearer explanation and proposed changes.

Doc Number: Korea JTC1/SC2 K1637 (2008.03.19)

Subject: Comments accompanying the letter ballot by Korea
  RE: SC2 N3982 (ISO/IEC 10646:2003 FPDAM5)

_XX_ DISAPPROVAL OF THE DRAFT FOR REASONS ON THE ATTACHED
    _XX_ Acceptance of these reasons and appropriate changes in the text
         will change our vote to approval

Technical comments.

T1. p. 1, right column

T1.1 Change as follows:

    1) (syllable-initial or initial consonant)
    --> (syllable-initial character or initial consonant)

    2) (syllable-peak or medial vowel)
    --> (syllable-peak character or medial vowel)

    3) (syllable-final or final consonant)
    --> (syllable-final character or final consonant)

T1.2 Rationale:
    1) Since "syllable-initial (-final)" implies consonantal,
       - "syllable-initial character" seems better than "syllable-initial
consonant".

    2) Likewise, since "syllable-peak" implies vowel, "syllable-peak
character" seems better than "syllable-peak vowel".

T2. p. 12,

T2.1 Replace the glyph for U110D with a correct one.

T2.2 Rationale:
    - The glyph for U110D is the same as that for U115F.
    - The glyph for U115F is fine.
    - However, the glyph for U110D is incorrect and therefore should be
changed.
    - We will provide the corrected font.

T3. Change six (6) phrases (three alternatives are given below).

T3.1 six (6) relevant phrases
   p. 13, left column, "Initial consonants"
   p. 14, left column, "Medial vowels"
   p. 14, right column, "Final consonants"
   p. 24, right column, "Initial consonants"
   p. 28, left column, "Medial vowels"
   p. 28, left column, "Final consonants"

T3.2 Rationale:
   - Our national position has been to use either "Choseong (Jungseong,
Jongseong)" or "syllable-initial (-peak, -final) character", as used in
ISO/IEC 10646 since the 1993 edition.

   - We suggest that one of the following alternatives be used:

   1) alternative 1: delete those six (6) phrases.
      We have not had those phrases in ISO/IEC 10646 since the 1993 edition.

   2) if we want to have some phrases in those 6 places, change as follows:

   2-1) alternative 2A:
   p. 13, left column, "Initial consonants"-> "Syllable-initial Characters"
   p. 14, left column, "Medial vowels"   -> "Syllable-peak Characters"
   p. 14, right column, "Final consonants"   -> "Syllable-final Characters"
   p. 24, right column, "Initial consonants"-> "Syllable-initial Characters"
   p. 28, left column, "Medial vowels"   -> "Syllable-peak Characters"
   p. 28, left column, "Final consonants"   -> "Syllable-final Characters"

   2-2) alternative 2B:
   p. 13, left column, "Initial consonants"
      -> "Syllable-initial Characters or Initial consonants"
   p. 14, left column, "Medial vowels"
      -> "Syllable-peak Characters or Medial vowels"
   p. 14, right column, "Final consonants"
      -> "Syllable-final Characters or Final consonants"
   p. 24, right column, "Initial consonants"
      -> "Syllable-initial Characters or Initial consonants"
   p. 28, left column, "Medial vowels"
      -> "Syllable-peak Characters or Medial vowels"
   p. 28, left column, "Final consonants"
      -> "Syllable-final Characters or Final consonants"

T4. Add an annotation for U+3180

   - According to Resolution M50.34 (Hangul Jamo additions), WG2 accepted to
add five annotations proposed in document WG2 N3172 (2006-09-27). However,
one of those five annotations is missing in the FPDAM5.
   - Therefore we request that one annotation be added as shown below:

   U+3180 HANGUL LETTER SSANGIEUNG
        => HANGUL LETTER SSANGIEUNG (ssangyesieung)

   - By the way, the relevant page (page 227 in ISO/IEC 10646:2003) is not
included in FPDAM5.

T5. p. 34, left column
   - For UF9B8, the correct mapping seems U96B7, NOT U96B8.
   - We suggest that
     1) delete 96B8 entry and
     2) keep 96B7 entry only.

T6. Change K5 entries in CJKU_SR.txt file

   - We decided to change the K5 reference format from "K5Hddddd" (decimal)
to "K5-hhhh" (hexadecimal) [see p. 2, right column of FPDAM5].
   - The change is well reflected in the code table of FPDAM5.
   - However, the change is NOT reflected in CJKU_SR.txt file of FPDAM5.
   - Therefore, we request that the change be reflected in CJKU_SR.txt file.

   - An example is shown below:

2A710;;;;K5H00418;;;;;
-->
2A710;;;;K5-01A2;;;;;

T7. Delete one line from CJKU_SR.txt file.

   - According to M51.11.g, we decided to delete one Hanja suggested by Rep.
of Korea: its reference is K5H00029 (or K5-001D in a new format).
   - CJKU_SR.txt file still contains that character.
   - We suggest that the following line be removed from CJKU_SR.txt file:

2B265;;;;K5H00029;;;;;

* * *

Dear SC2 Sec,

The UK vote on ISO/IEC 10646:2003 FPDAM5 (SC2 N3982), due on 2008-03-20>>

The UK votes to DISAPPROVE the amendment, with the following technical comments. Satisfactory resolution of comments T.1 and T.2 will change our vote to APPROVAL.

Technical Comments

T.1 ARIB Unified and Compatibility Ideographs

The UK has grave concerns over the encoding of the two unified ideographs 9FC4..9FC5 and the four compatibility ideographs FA6B..FA6E requested by Japan.

In particular, the UK objects to the encoding of U+9FC4 (ARIB#47) as a unified ideograph as it is a unifiable variant of U+6881. We note that this character has the perhaps unique distinction of having been removed from Amd.5 with one hand (it was U+2ACAD in PDAM5) and added back to Amd.5 with the other (see Resolutions M51.10-11).

On the other hand U+FA6D (ARIB#93) is not cognate with U+7953 according to the Chinese source, and so it may be a candidate for encoding as a unified ideograph.

There is also the wider issue of whether any new compatibility ideographs should be accepted for encoding now that the Ideographic Variation Database (IVD) is ready to be incorporated into the standard. We do not believe that it is helpful to have two different mechanisms for encoding unifiable ideographs (either as compatibility ideographs or by means of IVS sequences), as it can only cause confusion amongst users as to which mechanism is the appropriate one to use. With regard to the six ARIB characters, we note that five of them could be represented by means of the Adobe-Japan1 IVS collection <http://www.unicode.org/ivd/data/2007-12-14/>:

U+FA6B (ARIB#39) = <U+6075 U+E0101>
U+FA6C (ARIB#67) = <U+242EE U+E0101>
U+FA6D (ARIB#93) = <U+7953 U+E0101>
U+FA6E (ARIB#105) = <U+8218 U+E0101>
U+9FC4 (ARIB#47) = <U+6881 U+E0101>

We do not believe that new compatibility ideographs should be encoded when they could be represented as IVSes, which have the advantage over compatibility ideographs of not being lost during the process of normalization. However, it is not clear to us whether it would be appropriate to represent these ARIB characters by means of the corresponding Adobe-Japan1 IVSes or whether a new ARIB IVS collection should be defined instead. We hope that the issue of the quality control of the IVD, and the relationship between the IVD and ISO/IEC 10646 will be discussed at WG2 #52.

In the light of these concerns, and given that there has not been unanimous agreement on the encoding of these six characters at either WG2 #51 or IRG #29, we request that 9FC4..9FC5 and FA6B..FA6E be removed from Amd.5.

T.2 Tai Tham

The UK supports the recommendations given in the Tai Tham Ad-hoc Meeting Report (N3379).

We also request the addition of the two subjoined characters proposed in N3384 to Amd.5.

Title: Comments accompanying the US negative vote on FPDAM5 to ISO/IEC 10646:2003
Source: INCITS/L2
Action: Forward to INCITS

The US National Body is voting *no* with technical and editorial comments on the following SC2 ballot:

> SC2 N3982: ISO/IEC 10646: 2003/FPDAM 5, Information technology - Universal Multiple-Octet Coded Character Set (UCS) - AMENDMENT 5: Tai Tham, Tai Viet, Avestan, Egyptian Hieroglyphs, CJK Unified Ideographs Extension C, and other characters

Acceptance of technical comments T3-T8 will change our vote to *yes.*

T1. U Source reference

The amendment currently adds the Unicode U source "UTC" with the reference "The Unicode Standard 5.1-2008." We recommend to replace this reference by "Unicode Technical Report #45, U-source Ideographs", as this new technical report provides more details about the origin of those characters.

T2. Ideographic Variation Database

The end of the third paragraph of Clause 20.5 and the following note currently read:

> Variations sequences composed of a unified ideograph as the base character and one of VARIATION SELECTOR-17 to VARIATION SELECTOR-256 from the Supplementary Special-purpose Plane (SSP) are registered in the Ideographic Variation Database defined by Unicode Technical Standard #37.
>
>> NOTE 2 - The Ideographic Variation Database is currently empty. When entries are registered, these variation sequences will be referenced by this standard.

Following the procedure defined by Unicode Technical Standard #37, a new version of the Ideographic Variation Database has been accepted on December 12, 2007. This version is identified as '2007-12-14', and contains 14,651 sequences, covering the repertoire of the Adobe Japan1 reper-

toire. We suggest to replace NOTE 2 by the following paragraph or similar text:

> This version of the standard incorporates by reference the variation sequences listed in version 2007-12-14 of the Ideographic Variation Database, as described at <http://www.unicode.org/ivd/data/2007-12-14>.

T3. CJK Compatibility ideographs

While we entirely agree with the goal of establishing round-tripping between the ARIB character set and ISO/IEC 10646, we believe that there is now a better solution than encoding compability ideographs.

Consider the case of 恵 and 惠 which are distinct in the ARIB character set. In the model of ISO/IEC 10646, those two forms are unified as U+6075. To achieve round-tripping, the two forms must be mapped to different sequences of ISO/IEC 10646 characters.

The usual solution is to encode a compatibility ideograph, U+FA6B in this case, and to establish a canonical decomposition of that compability ideograph into the unified form, U+6075 in this case. The purpose of the decomposition is to account for the unification. Under that solution 恵 is mapped to the sequence <U+6075>, and 惠 is mapped to the sequence <U+FA6B>. However, this approach imposes a very severe constraint on implementations, as they can never normalize data; any normalization transforms <U+FA6B> into <U+6075>
and prevents round-tripping. Essentially, the canonical decomposition defeats the purpose of the compatibility ideograph.

With the advent of variation sequences, we have a better solution at our disposal. Indeed the variation sequence <U+6075, U+E0100> is targetting the form 恵 and the variation sequence <U+6075, U+E0101> is targetting the form 惠, so the ARIB characters can be mapped to those sequences and support round-tripping. Unlike the sequences of the usual solution, these variation sequences remain unchanged by normalization. This gives a much greater freedom to implementations.

The Ideographic Variation Database also contains sequences for the other three pairs of ARIB characters which are unified in ISO/IEC 10646.

In conclusion, we believe that the proposed characters U+FA6B..FA6E fail to effectively achieve the goal of round-tripping the ARIB character set, and that this goal can be achieved today using variation sequences already in the Ideographic Variation Database. We propose to not encode those four characters.

T4. Conflicting sources

The editor's note at the bottom of page 2 mentions that there is unresolved conflicting information

concerning KangXi source references. We would like these conflicts to be resolved before further progression of the Amendment.

T5. Names of Hangul jamos

The names of the three characters 11FD, A96E and A973 should have an additional "H" at their end. The correct names are:

11FD HANGUL JONGSEONG KIYEOK-KHIEUKH
A96E HANGUL CHOSEONG RIEUL-KHIEUKH
A973 HANGUL CHOSEONG PIEUP-KHIEUKH

(Apparently, the original version of the proposal WG2 N3168 had incorrect names, which in turn led to incorrect names in WG2 N3242, which is what was accepted by motion M50.34. As described in WG2 N3257, a revision of WG2 N3168 included the correct names as above.)

T6. Avestan separation point

The US NB remains opposed to the encoding of yet another middle dot punctuation at position 10B38 (AVESTAN SEPARATION POINT).

T7. Archaic Sinhala numerals

The US NB has received information that indicates that more investigation is needed for the Sinhala archaic digits and numbers (0DE7-0DEF and 0DF5-0DFF). The US NB would like those characters to be moved to a future amendment

T8. Tai Tham

The US NB supports the recommendations of the Tai Tham ad-hoc meeting as documented in WG2 N3379, as well as the inclusion of the two additional characters requested in WG2 N3384.

E1. Incorrect U Source header

On page 5, the header of the additional code chart fragment for the new characters 9FC4 and 9FC5 is "U Unicode". Those characters only have a J source, so the header should be "J".
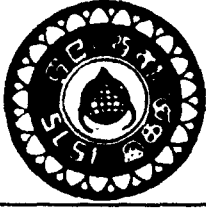
ශ්‍රී ලංකා ප්‍රමිති ආයතනය
இலங்கை கட்டளைகள் நிறுவனம்
SRI LANKA STANDARDS INSTITUTION

අංක 17, වික්ටෝරියා පෙදෙස, ඇල්විටිගල මාවත, කොළඹ 8.
இல. 17, விக்டோரியா பிளேஸ், எல்விட்டிகல மாவத்தை, கொழும்பு 8.
No. 17, VICTORIA PLACE, ELVITIGALA MAWATHA, COLOMBO 8, SRI LANKA.

පැ.පෙ.
த. பெ.
P. O. Box } 17

දුරකථනය
தொலைபேசி
TELEPHONE } 94 11 2671567-72

ෆැක්ස්
பாக்ஸ்
FAX } 94 11 2671579 / 94 11 2672617

විද්‍යුත් තැපෑල: 'ප්‍රමික', කොළඹ
தந்தி: 'பிரமிகா', கொழும்பு
Telegrams: 'pramika', Colombo

E-mail: slsi@slsi.slt.lk

Web Site: http://www.slsi.lk

ඔබේ යොමුව
உமது குறிப்பு
Your Ref. }

අපේ යොමුව
எமது குறிப்பு
Our Ref. }

දිනය
திகதி
Date }

SC/20

2008-01-11

Ms. Toshiko Kimura, Secretary
ISO/IEC JTC 1 SC2 Secretariat
IPSJ/ITSCJ
Room 308-3, Kikai-Shinko-Kaikan Bldg., 3-5-8,
Shiba-Koen, Minato-Ku
Tokyo 105-0011
Japan

Dear Madam,

SINHALA CHARACTER CODE PAGE - ISO/IEC 10646 : 2003 /FPD Amd 5

Sri Lanka Standards Institution (SLSI) is the national Standards body in Sri Lanka, and is also a member of ISO and IEC representing Sri Lanka. SLSI established the national standard for Sinhala Character Code for Information Interchange (SLS 1134), initially in year 1989. This Sri Lanka standard was revised later in year 1996 and year 2004 in association with the Information and Communication Technology Agency of Sri Lanka (ICTA) who is the government body responsible for promoting the use of ICT in Sri Lanka including ICT in Sri Lanka's local languages Sinhala and Tamil. Sinhala Character Code Set in SLS 1134, which is also in harmony with the Unicode, was presented to ISO for inclusion in ISO/IEC 10646 by SLSI.

SLSI recently found that a paper has been submitted by Mr. Michael Everson for inclusion of archaic Sinhala numerals in the Sinhala Code range, and that JTC 1 SC2 has accepted and registered as ISO/IEC 10646 : 2003 /FPD Amd 5 which is being processed to issue an amendment to Sinhala Character Code range.
Sinhala numerals had been used in the first and the second centuries AD, and later there had been different numerals used in the fifteenth century, all these have been replaced by Arabic numerals (European), which are used today.

Page 01 of 03

සභාපති
தலைவர்
Chairman 94 11 2671573

අධ්‍යක්ෂ ජනරාල්/ප්‍රධාන විධායක
பணிப்பாளர் நாயகம்/பிரதம நிறைவேற்று
Director General/Chief Executive 94 11 2671574

Sri Lanka is the only country where Sinhala is used and Sinhala has been declared as a National and Official language. Any amendment to this Character Code range should be in harmony with the actual use of Sinhala and attention should be given to Sinhala used in Sri Lanka. Since it is one of Sri Lanka's national and official languages, and the majority of Sri Lankans use Sinhala, the Sri Lanka Standard Institution, is responsible for ensuring that the ISO Code range is in line with contemporary Sinhala. Any proposal for changes to the existing Sinhala Character Code range should be communicated to the National Standards Body (SLSI) of Sri Lanka, as Sri Lanka is the only country where majority of people communicate in Sinhala.

We wish to conduct a careful study into this issue to prevent misrepresentation of our National and Official Language. However, a preliminary investigation by ICTA and the University of Colombo School of Computing (UCSC) shows that while Mr. Everson addresses only ONE type of Sinhala Numerals, but there are in fact at least five types that had been used in our past of history over 2000 years. Out of these five types at least three are still in use in very specialized areas which is certainly not known to the average users of Sinhala. We wish to therefore raise the following observations:

1. Sinhala numbers based on Brahmi letters have been used around the first century to third century AD.
2. The form of numerals known as "Lith Anka" (calendar numerals) and had been used in the 18th and early 19th centuries. These seem to be what have been proposed by Mr. Everson.
3. A method of writing numbers in Sinhala called "Katapaydia" has been used in horoscopes, etc. and is still in use in the 21st Century, and is well known in the astrological community. This uses Sinhala letters to represent numbers. This appears to be the most common method (other than European numerals) for representing numbers in Sinhala.
4. Another form of writing numbers in Sinhala is described in Dr. Hendrick de Silva Hettigoda's famous astrology book written in 1967, which is another type of "Lith Anka", based on modified Sinhala letters and symbols. This method too has been used in recent times.
5. Ola-leaf manuscripts are numbered using the full Sinhala alphabet in order.
6. All Official, Legal, and Private documents use Arabic numerals and Sinhala numerals are currently not in common use.

In view of the above points, it is our considered view that there is currently no agreement on what constitutes Sinhala Numerals and how they should be represented in UCS. It is also clear that there is no pressing or urgent need to immediately encode Sinhala Numerals. We do not agree with the justification provided in Mr. Everson's proposal that contact been made to members of the user community, as the Institutions such as SLSI, ICTA and the Department Official Languages of Sri Lanka were not even made aware of the proposal, and any such urgent need has not arisen for this amendment.

It is our request that the above referenced proposal N3195R be deferred so that there is sufficient time for the relevant stakeholders to carefully study the above points, to have a sound discussion on the subject amongst a wider audience. This process would definitely help to develop an official proposal from Sri Lanka through SLSI. We further kindly request that ISO refrains from taking any further action in this regard without our concurrence.

Thank you.

Yours faithfully,

Dr. L N Senaweera
Actg. Director General
SRI LANKA STANDARDS INSTITUTION

Copy to : ISO Secretary General