

Universal Multiple-Octet Coded Character Set
 International Organization for Standardization
 Organisation Internationale de Normalisation
 Международная организация по стандартизации

Doc Type: Working Group Document

Title: Proposal to add archaic numbers for Sinhala to the BMP and SMP of the UCS

Source: Ireland

Status: National Body Contribution

Action: For consideration by JTC1/SC2/WG2 and UTC

Replaces: N1473R, N3195R

Date: 2010-08-08

Present-day Sri Lankans use the European digits 0123456789 exclusively, and the older numbering systems are not known by most people literate in Sinhala. N1473R, dated 1997-03-01, proposed what is now known to be an incomplete set of these numbers. UTR#2 published in 1992, contained a proposal for Sinhala written by Andy Daniels which included the digits as listed above as well as NUMBER ONE HUNDRED and NUMBER ONE THOUSAND (but not the tens). The numbers were not considered well-attested and they were removed from the proposal for further study as their encoding was not urgent. In February 2007, document N3195R requested twenty additional characters to be added to the UCS. Experts from Sri Lanka requested time for further study, and recently a book was published (Wijayawardhana 2009) in which the results of that study indicate that the numbers requested in N3195R are justified for encoding, but that an additional set of numbers should also be encoded. A more modern set, called *lith illakkum* in Sinhala were used in ephemerides for casting horoscopes and making astrological calculations. The older, pre-decimal set given in N3195R, was previously approved for encoding by WG2 and the UTC.

The Archaic Sinhala number system, which seems not to have been used after 1815, was additive. Because there was no zero, a separate character was used for each of the tens. The usage of the numbers is illustrated here:

	෧ 1	෨ 2	෩ 3	෪ 4	෫ 5	෬ 6	෭ 7	෮ 8	෯ 9
෧෦ 10	෧෧ 11	෧෨ 12	෧෩ 13	෧෪ 14	෧෫ 15	෧෬ 16	෧෭ 17	෧෮ 18	෧෯ 19
෨෦ 20	෨෧ 21	෨෨ 22	෨෩ 23	෨෪ 24	෨෫ 25	෨෬ 26	෨෭ 27	෨෮ 28	෨෯ 29
෩෦ 30	෩෧ 31	෩෨ 32	෩෩ 33	෩෪ 34	෩෫ 35	෩෬ 36	෩෭ 37	෩෮ 38	෩෯ 39
෪෦ 40	෪෧ 41	෪෨ 42	෪෩ 43	෪෪ 44	෪෫ 45	෪෬ 46	෪෭ 47	෪෮ 48	෪෯ 49
෫෦ 50	෫෧ 51	෫෨ 52	෫෩ 53	෫෪ 54	෫෫ 55	෫෬ 56	෫෭ 57	෫෮ 58	෫෯ 59
෬෦ 60	෬෧ 61	෬෨ 62	෬෩ 63	෬෪ 64	෬෫ 65	෬෬ 66	෬෭ 67	෬෮ 68	෬෯ 69
෭෦ 70	෭෧ 71	෭෨ 72	෭෩ 73	෭෪ 74	෭෫ 75	෭෬ 76	෭෭ 77	෭෮ 78	෭෯ 79
෮෦ 80	෮෧ 81	෮෨ 82	෮෩ 83	෮෪ 84	෮෫ 85	෮෬ 86	෮෭ 87	෮෮ 88	෮෯ 89
෯෦ 90	෯෧ 91	෯෨ 92	෯෩ 93	෯෪ 94	෯෫ 95	෯෬ 96	෯෭ 97	෯෮ 98	෯෯ 99
෧෦෦ 100	෧෦෧ 101	෧෦෨ 102	෧෦෩ 103	෧෦෪ 104	෧෦෫ 105	෧෦෬ 106	෧෦෭ 107	෧෦෮ 108	෧෦෯ 109
෧෦෧෦ 111	෧෦෧෧ 222		෧෦෧෨ 333		෧෦෧෩ 444		෧෦෧෪ 555		
෧෦෧෧෦ 666	෧෦෧෧෧ 777		෧෦෧෧෨ 888		෧෦෧෧෩ 999				
෧෦෧෧෧෦ 1000	෧෦෧෧෧෧෦ 1963		෧෦෧෧෧෧෨ 2007		෧෦෧෧෧෧෩ 2010				

The astrological Sinhala number system appears to have been in use even up to the earlier part of the 20th century, is a true decimal system, including the use of the digit zero. The usage of the numbers is illustrated here:

ඉ 0	෧ 1	෨ 2	෩ 3	෪ 4	෫ 5	෬ 6	෭ 7	෮ 8	෯ 9
෧෦ 10	෧෧ 11	෧෨ 12	෧෩ 13	෧෪ 14	෧෫ 15	෧෬ 16	෧෭ 17	෧෮ 18	෧෯ 19
෨෦ 20	෨෧ 21	෨෨ 22	෨෩ 23	෨෪ 24	෨෫ 25	෨෬ 26	෨෭ 27	෨෮ 28	෨෯ 29
෩෦ 30	෩෧ 31	෩෨ 32	෩෩ 33	෩෪ 34	෩෫ 35	෩෬ 36	෩෭ 37	෩෮ 38	෩෯ 39
෪෦ 40	෪෧ 41	෪෨ 42	෪෩ 43	෪෪ 44	෪෫ 45	෪෬ 46	෪෭ 47	෪෮ 48	෪෯ 49
෫෦ 50	෫෧ 51	෫෨ 52	෫෩ 53	෫෪ 54	෫෫ 55	෫෬ 56	෫෭ 57	෫෮ 58	෫෯ 59
෬෦ 60	෬෧ 61	෬෨ 62	෬෩ 63	෬෪ 64	෬෫ 65	෬෬ 66	෬෭ 67	෬෮ 68	෬෯ 69
෭෦ 70	෭෧ 71	෭෨ 72	෭෩ 73	෭෪ 74	෭෫ 75	෭෬ 76	෭෭ 77	෭෮ 78	෭෯ 79
෮෦ 80	෮෧ 81	෮෨ 82	෮෩ 83	෮෪ 84	෮෫ 85	෮෬ 86	෮෭ 87	෮෮ 88	෮෯ 89
෯෦ 90	෯෧ 91	෯෨ 92	෯෩ 93	෯෪ 94	෯෫ 95	෯෬ 96	෯෭ 97	෯෮ 98	෯෯ 99
෧෦෦ 100	෧෦෧ 101	෧෦෨ 102	෧෦෩ 103	෧෦෪ 104	෧෦෫ 105	෧෦෬ 106	෧෦෭ 107	෧෦෮ 108	෧෦෯ 109
෧෧෧ 111		෧෧෨ 222		෧෧෩ 333		෧෧෪ 444		෧෧෫ 555	
෧෧෧ 666		෧෧෭ 777		෧෧෮ 888		෧෧෯ 999			
෧෦෦෦ 1000		෧෧෧෩ 1963		෧෦෦෭ 2007		෧෦෦෦ 2010			

Unicode Character Properties

```

0DE6;SINHALA DIGIT ONE;No;0;L;;;0;N;;;;;
0DE7;SINHALA DIGIT ONE;No;0;L;;;1;N;;;;;
0DE8;SINHALA DIGIT TWO;No;0;L;;;2;N;;;;;
0DE9;SINHALA DIGIT THREE;No;0;L;;;3;N;;;;;
0DEA;SINHALA DIGIT FOUR;No;0;L;;;4;N;;;;;
0DEB;SINHALA DIGIT FIVE;No;0;L;;;5;N;;;;;
0DEC;SINHALA DIGIT SIX;No;0;L;;;6;N;;;;;
0DED;SINHALA DIGIT SEVEN;No;0;L;;;7;N;;;;;
0DEE;SINHALA DIGIT EIGHT;No;0;L;;;8;N;;;;;
0DEF;SINHALA DIGIT NINE;No;0;L;;;9;N;;;;;

111E1;SINHALA DIGIT ONE;No;0;L;;;1;N;;;;;
111E2;SINHALA DIGIT TWO;No;0;L;;;2;N;;;;;
111E3;SINHALA DIGIT THREE;No;0;L;;;3;N;;;;;
111E4;SINHALA DIGIT FOUR;No;0;L;;;4;N;;;;;
111E5;SINHALA DIGIT FIVE;No;0;L;;;5;N;;;;;
111E6;SINHALA DIGIT SIX;No;0;L;;;6;N;;;;;
111E7;SINHALA DIGIT SEVEN;No;0;L;;;7;N;;;;;
111E8;SINHALA DIGIT EIGHT;No;0;L;;;8;N;;;;;
111E9;SINHALA DIGIT NINE;No;0;L;;;9;N;;;;;
111EA;SINHALA NUMBER TEN;No;0;L;;;10;N;;;;;
111EB;SINHALA NUMBER TWENTY;No;0;L;;;20;N;;;;;
111EC;SINHALA NUMBER THIRTY;No;0;L;;;30;N;;;;;
111ED;SINHALA NUMBER FORTY;No;0;L;;;40;N;;;;;
111EE;SINHALA NUMBER FIFTY;No;0;L;;;50;N;;;;;
111EF;SINHALA NUMBER SIXTY;No;0;L;;;60;N;;;;;
111F0;SINHALA NUMBER SEVENTY;No;0;L;;;70;N;;;;;
111F1;SINHALA NUMBER EIGHTY;No;0;L;;;80;N;;;;;
111F2;SINHALA NUMBER NINETY;No;0;L;;;90;N;;;;;
111F3;SINHALA NUMBER ONE HUNDRED;No;0;L;;;100;N;;;;;
111F4;SINHALA NUMBER ONE THOUSAND;No;0;L;;;1000;N;;;;;

```

Bibliography

Wijayawardhana, Harsha. 2009. *Numerations in the Sinhala Language*. Edited by Aruni Goonetilleke. Colombo: Strategic Communications and Media Unit - ICTA. ISBN 978-955-1199-05-0

	0D8	0D9	0DA	0DB	0DC	0DD	0DE	0DF
0		ඌ 0D90	ඬ 0DA0	ඨ 0DB0	ඪ 0DC0	ණ 0DD0		
1		ඬ 0D91	ණ 0DA1	ඪ 0DB1	ණ 0DC1	ඬ 0DD1		
2	ං 0D82	ඌ 0D92	ණ 0DA2		ඪ 0DC2	ඬ 0DD2		෦෦෦ 0DF2
3	ං 0D83	ඌ 0D93	ණ 0DA3	ඨ 0DB3	ඪ 0DC3	ඬ 0DD3		ඌ 0DF3
4		ඬ 0D94	ණ 0DA4	ඪ 0DB4	ඪ 0DC4	ඬ 0DD4		ණ 0DF4
5	ණ 0D85	ඬ 0D95	ණ 0DA5	ඪ 0DB5	ඪ 0DC5			
6	ණ 0D86	ඬ 0D96	ණ 0DA6	ඪ 0DB6	ඪ 0DC6	ඬ 0DD6	ණ 0DE6	
7	ණ 0D87		ඪ 0DA7	ඪ 0DB7			ණ 0DE7	
8	ණ 0D88		ඪ 0DA8	ඪ 0DB8		ණ 0DD8	ණ 0DE8	
9	ණ 0D89		ඪ 0DA9	ඪ 0DB9		ණ 0DD9	ණ 0DE9	
A	ණ 0D8A	ණ 0D9A	ණ 0DA A	ඪ 0DBA	ඪ 0DCA	ණ 0DDA	ණ 0DEA	
B	ණ 0D8B	ඬ 0D9B	ණ 0DAB	ඪ 0DBB		ණ 0DDB	ණ 0DEB	
C	ණ 0D8C	ඬ 0D9C	ණ 0DAC			ණ 0DDC	ණ 0DEC	
D	ණ 0D8D	ඬ 0D9D	ණ 0DAD	ඪ 0DBD		ණ 0DDD	ණ 0DED	
E	ණ 0D8E	ඬ 0D9E	ණ 0DAE			ණ 0DDE	ණ 0DEE	
F	ණ 0D8F	ඬ 0D9F	ණ 0DAF		ණ 0DCF	ණ 0DDF	ණ 0DEF	

Various signs





- OD82 ෧෧ SINHALA SIGN ANUSVARAYA
= anusvara
- OD83 ෧෧ SINHALA SIGN VISARGAYA
= visarga

Independent vowels

- OD85 අ SINHALA LETTER AYANNA
= sinhala letter a
- OD86 ආ SINHALA LETTER AAYANNA
= sinhala letter aa
- OD87 ඈ SINHALA LETTER AEYANNA
= sinhala letter ae
- OD88 ඈ SINHALA LETTER AEEYANNA
= sinhala letter aae
- OD89 ඉ SINHALA LETTER IYANNA
= sinhala letter i
- OD8A ඊ SINHALA LETTER IYANNA
= sinhala letter ii
- OD8B උ SINHALA LETTER UYANNA
= sinhala letter u
- OD8C උෟ SINHALA LETTER UUYANNA
= sinhala letter uu
- OD8D ඊෟ SINHALA LETTER IRUYANNA
= sinhala letter vocalic r
- OD8E ඊෟෟ SINHALA LETTER IRUUYANNA
= sinhala letter vocalic rr
- OD8F ඊ SINHALA LETTER ILUYANNA
= sinhala letter vocalic l
- OD90 ඊෟ SINHALA LETTER ILUUYANNA
= sinhala letter vocalic ll
- OD91 ඊ SINHALA LETTER EYANNA
= sinhala letter e
- OD92 ඊ SINHALA LETTER EEYANNA
= sinhala letter ee
- OD93 ඊ෧ SINHALA LETTER AIYANNA
= sinhala letter ai
- OD94 ඊ෧ SINHALA LETTER OYANNA
= sinhala letter o
- OD95 ඊ෧ SINHALA LETTER OUYANNA
= sinhala letter oo
- OD96 ඊ෧ෟ SINHALA LETTER AUYANNA
= sinhala letter au

Consonants

- OD9A ක SINHALA LETTER ALPAPRAANA KAYANNA
= sinhala letter ka
- OD9B ක SINHALA LETTER MAHAAPRAANA KAYANNA
= sinhala letter kha
- OD9C ග SINHALA LETTER ALPAPRAANA GAYANNA
= sinhala letter ga
- OD9D ග SINHALA LETTER MAHAAPRAANA GAYANNA
= sinhala letter gha
- OD9E ක SINHALA LETTER KANTAJA NAASIKYAYA
= sinhala letter nga
- OD9F න SINHALA LETTER SANYAKA GAYANNA
= sinhala letter nnga
- ODA0 ට SINHALA LETTER ALPAPRAANA CAYANNA
= sinhala letter ca
- ODA1 ජ SINHALA LETTER MAHAAPRAANA CAYANNA
= sinhala letter cha
- ODA2 ජ SINHALA LETTER ALPAPRAANA JAYANNA
= sinhala letter ja
- ODA3 ක් SINHALA LETTER MAHAAPRAANA JAYANNA
= sinhala letter jha

- ODA4 ක් SINHALA LETTER TAALUJA NAASIKYAYA
= sinhala letter nya
- ODA5 ක් SINHALA LETTER TAALUJA SANYOOGA
NAASIKYAYA
= sinhala letter jnya
- ODA6 ජ SINHALA LETTER SANYAKA JAYANNA
= sinhala letter nyja
- ODA7 ට SINHALA LETTER ALPAPRAANA TTAYANNA
= sinhala letter tta
- ODA8 ට SINHALA LETTER MAHAAPRAANA TTAYANNA
= sinhala letter ttha
- ODA9 ට SINHALA LETTER ALPAPRAANA DDAYANNA
= sinhala letter dda
- ODAA ට SINHALA LETTER MAHAAPRAANA DDAYANNA
= sinhala letter ddha
- ODAB න් SINHALA LETTER MUURDHAJA NAYANNA
= sinhala letter nna
- ODAC ට SINHALA LETTER SANYAKA DDAYANNA
= sinhala letter nnda
- ODAD න SINHALA LETTER ALPAPRAANA TAYANNA
= sinhala letter ta
- ODAE ට SINHALA LETTER MAHAAPRAANA TAYANNA
= sinhala letter tha
- ODAF ට SINHALA LETTER ALPAPRAANA DAYANNA
= sinhala letter da
- ODB0 ට SINHALA LETTER MAHAAPRAANA DAYANNA
= sinhala letter dha
- ODB1 න SINHALA LETTER DANTAJA NAYANNA
= sinhala letter na
- ODB2  <reserved>
- ODB3 ට SINHALA LETTER SANYAKA DAYANNA
= sinhala letter nda
- ODB4 ප SINHALA LETTER ALPAPRAANA PAYANNA
= sinhala letter pa
- ODB5 ප SINHALA LETTER MAHAAPRAANA PAYANNA
= sinhala letter pha
- ODB6 බ SINHALA LETTER ALPAPRAANA BAYANNA
= sinhala letter ba
- ODB7 බ SINHALA LETTER MAHAAPRAANA BAYANNA
= sinhala letter bha
- ODB8 ම SINHALA LETTER MAYANNA
= sinhala letter ma
- ODB9 ම SINHALA LETTER AMBA BAYANNA
= sinhala letter mba
- ODBA ය SINHALA LETTER YAYANNA
= sinhala letter ya
- ODBB ර SINHALA LETTER RAYANNA
= sinhala letter ra
- ODBC  <reserved>
- ODBD ට SINHALA LETTER DANTAJA LAYANNA
= sinhala letter la
• dental
- ODBE  <reserved>
- ODBF  <reserved>
- ODC0 ට SINHALA LETTER VAYANNA
= sinhala letter va
- ODC1 ො SINHALA LETTER TAALUJA SAYANNA
= sinhala letter sha
- ODC2 ො SINHALA LETTER MUURDHAJA SAYANNA
= sinhala letter ssa
• retroflex
- ODC3 ස SINHALA LETTER DANTAJA SAYANNA
= sinhala letter sa
• dental
- ODC4 හ SINHALA LETTER HAYANNA
= sinhala letter ha

- 0DC5 ට SINHALA LETTER MUURDHAJA LAYANNA
= sinhala letter lla
• retroflex
- 0DC6 ට SINHALA LETTER FAYANNA
= sinhala letter fa

Sign

- 0DCA ට SINHALA SIGN AL-LAKUNA
= virama

Dependent vowel signs

- 0DCF ට SINHALA VOWEL SIGN AELA-PILLA
= sinhala vowel sign aa
- 0DD0 ට SINHALA VOWEL SIGN KETTI AEDA-PILLA
= sinhala vowel sign ae
- 0DD1 ට SINHALA VOWEL SIGN DIGA AEDA-PILLA
= sinhala vowel sign aae
- 0DD2 ට SINHALA VOWEL SIGN KETTI IS-PILLA
= sinhala vowel sign i
- 0DD3 ට SINHALA VOWEL SIGN DIGA IS-PILLA
= sinhala vowel sign ii
- 0DD4 ට SINHALA VOWEL SIGN KETTI PAA-PILLA
= sinhala vowel sign u
- 0DD5  <reserved>
- 0DD6 ට SINHALA VOWEL SIGN DIGA PAA-PILLA
= sinhala vowel sign uu
- 0DD7  <reserved>
- 0DD8 ට SINHALA VOWEL SIGN GAETTA-PILLA
= sinhala vowel sign vocalic r
- 0DD9 ට SINHALA VOWEL SIGN KOMBUVA
= sinhala vowel sign e
- 0DDA ට SINHALA VOWEL SIGN DIGA KOMBUVA
= sinhala vowel sign ee
≡ 0DD9 ට 0DCA ට
- 0DDB ට SINHALA VOWEL SIGN KOMBU DEKA
= sinhala vowel sign ai

Two-part dependent vowel signs

These vowel signs have glyph pieces which stand on both sides of the consonant; they follow the consonant in logical order, and should be handled as a unit for most processing.

- 0DDC ට SINHALA VOWEL SIGN KOMBUVA HAA AELA-PILLA
= sinhala vowel sign o
≡ 0DD9 ට 0DCF ට
- 0DDD ට SINHALA VOWEL SIGN KOMBUVA HAA DIGA AELA-PILLA
= sinhala vowel sign oo
≡ 0DDC ට 0DCA ට
- 0DDE ට SINHALA VOWEL SIGN KOMBUVA HAA GAYANUKITTA
= sinhala vowel sign au
≡ 0DD9 ට 0DDF ට

Dependent vowel sign

- 0DDF ට SINHALA VOWEL SIGN GAYANUKITTA
= sinhala vowel sign vocalic l

Historical digits

- 0DE6 ෦ SINHALA DIGIT ZERO
- 0DE7 ෧ SINHALA DIGIT ONE
- 0DE8 ෨ SINHALA DIGIT TWO
- 0DE9 ෩ SINHALA DIGIT THREE
- 0DEA ෪ SINHALA DIGIT FOUR
- 0DEB ෫ SINHALA DIGIT FIVE
- 0DEC ෬ SINHALA DIGIT SIX

0DED ෭ SINHALA DIGIT SEVEN

0DEE ෮ SINHALA DIGIT EIGHT













0DEF ෯ SINHALA DIGIT NINE

Additional dependent vowel signs

- 0DF2 ට SINHALA VOWEL SIGN DIGA GAETTA-PILLA
= sinhala vowel sign vocalic rr
- 0DF3 ට SINHALA VOWEL SIGN DIGA GAYANUKITTA
= sinhala vowel sign vocalic ll

Punctuation

- 0DF4  SINHALA PUNCTUATION KUNDDALIYA

	111E	111F
0		෧ 111F0
1	෨ 111E1	෩ 111F1
2	෪ 111E2	෫ 111F2
3	෬ 111E3	෭ 111F3
4	෮ 111E4	෯ 111F4
5	10 111E5	
6	11 111E6	
7	12 111E7	
8	13 111E8	
9	14 111E9	
A	15 111EA	
B	16 111EB	
C	17 111EC	
D	18 111ED	
E	19 111EE	
F	20 111EF	

Historical digits

These digits are not used with a zero

111E1	෧	SINHALA ARCHAIC DIGIT ONE
111E2	෨	SINHALA ARCHAIC DIGIT TWO
111E3	෪	SINHALA ARCHAIC DIGIT THREE
111E4	෬	SINHALA ARCHAIC DIGIT FOUR
111E5	෮	SINHALA ARCHAIC DIGIT FIVE
111E6	෯	SINHALA ARCHAIC DIGIT SIX
111E7	10	SINHALA ARCHAIC DIGIT SEVEN
111E8	11	SINHALA ARCHAIC DIGIT EIGHT
111E9	12	SINHALA ARCHAIC DIGIT NINE

Historical numbers

111EA	15	SINHALA ARCHAIC NUMBER TEN
111EB	16	SINHALA ARCHAIC NUMBER TWENTY
111EC	17	SINHALA ARCHAIC NUMBER THIRTY
111ED	18	SINHALA ARCHAIC NUMBER FORTY
111EE	19	SINHALA ARCHAIC NUMBER FIFTY
111EF	20	SINHALA ARCHAIC NUMBER SIXTY
111F0	෧෫	SINHALA ARCHAIC NUMBER SEVENTY
111F1	෧෬	SINHALA ARCHAIC NUMBER EIGHTY
111F2	෧෭	SINHALA ARCHAIC NUMBER NINETY
111F3	෧෮	SINHALA ARCHAIC NUMBER ONE HUNDRED
111F4	෧෯	SINHALA ARCHAIC NUMBER ONE THOUSAND

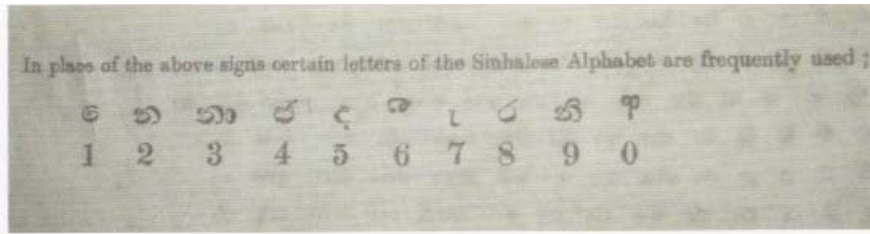


Figure 1. Sample from Wijayawardhana 2009 showing a sample by W. A. de Silva in which he has substituted letter glyphs for the decimal numbers.

දිනේ ඉලක්කම්

ට	ඟ	ඟා	ජ	ඳ
1	2	3	4	5
ආ	උ	ඌ	ඹ	ඹා
6	7	8	9	10
ටට	ටඟ	ටඟා	ජජ	ජඳ
11	12	13	14	15
ඹ	ඹඋ	ඹඌ	ඹඹ	ඹඹා
16	17	18	19	20
ඟට	ඟඟ	ඟඟා	ඟජ	ඟඳ
21	22	23	24	25

Figure 2. Sample from Wijayawardhana 2009 showing typographic forms of the decimal numbers in a document by Dr P. De S. Kularatne. As Wijayawardhana points out, the number 20 is incorrectly depicted.

A. Administrative

1. Title

Proposal to add archaic numbers for Sinhala to the BMP and SMP of the UCS

2. Requester's name

Ireland

3. Requester type (Member body/Liaison/Individual contribution)

National Body contribution.

4. Submission date

2010-08-08

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

Yes.

6b. More information will be provided later

No.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

No.

1b. Proposed name of script

1c. The proposal is for addition of character(s) to an existing block

Yes

1d. Name of the existing block

Sinhala

2. Number of characters in proposal

20

3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)

Category A.

4a. Is a repertoire including character names provided?

Yes.

4b. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?

Yes.

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

Yes.

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database <http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

Yes. See N1473R, N3195R

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

No.

2b. If YES, with whom?

2c. If YES, available relevant documents

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

The characters have historical use in Sinhala script.

4a. The context of use for the proposed characters (type of use; common or rare)

Used rarely today but found in documents prior to the 19th century CE.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

Publications describing the history of the script, ephemerides.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Only some of them.

6b. If YES, is a rationale provided?

No.

6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

No.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

No.

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?