2007-01-09

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

Doc Type:Working Group DocumentTitle:Proposal to add archaic numbers for Sinhala to the BMP of the UCSSource:Michael EversonStatus:Individual ContributionAction:For consideration by JTC1/SC2/WG2 and UTCReplaces:N1473RDate:2007-01-09

This document requests twenty additional characters to be added to the UCS and contains the proposal summary form. The characters are as follows:

**0DE7** SINHALA DIGIT ONE **0DE8** SINHALA DIGIT TWO **0DE9** SINHALA DIGIT THREE **0DEA** SINHALA DIGIT FOUR SINHALA DIGIT FIVE **0DEB 0DEC** SINHALA DIGIT SIX **0DED** SINHALA DIGIT SEVEN **ODEE** SINHALA DIGIT EIGHT **0DEF** SINHALA DIGIT NINE 0DF5 SINHALA NUMBER TEN 0DF6 SINHALA NUMBER TWENTY 0DF7 SINHALA NUMBER THIRTY **0DF8** SINHALA NUMBER FORTY 0DF9 SINHALA NUMBER FIFTY **0DFA** SINHALA NUMBER SIXTY **0DFB** SINHALA NUMBER SEVENTY **0DFC** SINHALA NUMBER EIGHTY 0DFD SINHALA NUMBER NINETY **0DFE** SINHALA NUMBER ONE HUNDRED **0DFF** SINHALA NUMBER ONE THOUSAND

**Background.** Present-day Sri Lankans use the European digits 0123456789 exclusively, and the older numbering system is not known by most people literate in Sinhala. N1473R, dated 1997-03-01, proposed the characters listed here. 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 October 1998, December 1999, October 2000, February 2004 and most recently in December 2006 people have raised the issue of Sinhala numbers either on the Unicode list or with me directly. At this point we have a larger number of citations for the numbers, and it seems appropriate to encode them now.

The old Sinhala number system 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:

	<b>N</b> 1	<b>へ</b> 2	<b>G</b> n 3	എ 4	On 5	<b>O</b> 6	ପ୍ରା <sub>7</sub>	<u>س</u> 8	<b>(9</b> )
ର୍ଦ୍ଦେ 10	<b>ର୍ଭୋତା</b> 11	ଭୋଇ 12	യിനെ 13	<b>പ്രേസ്ത</b> 14	<b>ଓ</b> ଧାତିଲ 15	ଭୋତି 16	<b>ର୍ଭ୍ୟର୍ମ</b> 17	ଭୋଫ୍ର 18	ଭାଙ୍ଫା 19
සි 20	<b>ଦ୍ଧିର</b> 21	ద్దం 22	En 23	සිෆු 24	Em 25	සිට 26	င်္ဆြဂူ 27	සිෆු 28	සි <b>ලා</b> 29
S 30	<b>ଟ୍ଟର</b> 31	Sa 32	Sn 33	<b>හි</b> කු 34	Son 35	SO 36	ଚ୍ଚିର୍ 37	පිෆු 38	S ( 39
<b>େ</b> ଅ 40	<b>େ</b> ୍କର 41	മ്പുര 42	ക്കേസം 43	ക്ഷപ്പ 44	ക്ഷം 45	ත සිට 46	ଦ୍ଧେମ୍ବା 47	ති අසු 48	<b>୍ଦେଖ୍ୟ</b> 49
ති 50	<b>ର୍ଦ୍ଧର</b> 51	ര് 52	ര്പ 53	ഹ്പ്പെ 54	Bm 55	ඟිට 56	င်္သေဂျ 57	ඟිෆු 58	ත්ල 59
ম্পে 60	ଙ୍କାର 61	ന്ദ്രം 62	ന്ദിന 63	ന്ദ്രന്നു 64	<b>ന്ദ്ര</b> സ 65	ଆଠି 66	ମ୍ବାର୍ମ୍ବ 67	ଔାର୍ଫ୍ର 68	ଙ୍କା 🕑 69
ମ୍ବରୋ 70	ମ୍ବର୍ଦ୍ଧାତା 71	രുക്രം 72	രുഗ്രേഹ 73	രുഗ്കന്നു 74	രുമി സെ 75	୯୧୪୦୦ 76	ଅବ୍ୟୁର୍ଦ୍ଧାମ୍ଲା 77	୧୧୪୦୦୦ 78	ଫ୍ରେଆ 79
B 80	<b>ଧିର</b> 81	En 82	En 83	පිෆු 84	Em 85	පිට 86	င်းရှု 87	පිლ 88	S ( 89
<b>B</b> 90	ଧିବ ୨۱	<b>පි</b> ი 92	Sn 93	හිආ 94	8 m 95	SO 96	ଦ୍ଧିର୍ ୨୨	හිෆු 98	<b>୍ଷେଳା</b> ୨୨
സ്റ്റി 100	ୁଆର 101	<b>സ്പി</b> പെ 102	<b>സ്റ്റി</b> റെ 103	<b>സ്റ്റിന്റ</b> 104	<b>സ്റ്റി</b> നെ 105	MIO 106	<b>സ്പിറ്റി</b> 107	സ്റ്റിസ് 108	<b>Million</b> 109
ଲ୍ଲାଜ୍ୟାଠ	111	୶ୣ୷୲ଘର	222	๛๛เย	u 333	എസ്വാത്ഷം	<b>ግ</b> 444	๛൬๗๛	n 555
୦୯୦୦	666	ମାର୍ଭାରଜ୍ୟୁ	Ŋ 777	෩෩෦ඁ෨ඁ	888	<u>୍</u> ମାକ୍ଷ୍ୟାର୍ଷ୍ଣ	999		
ണ്ണ 1000		~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	21 <b>C</b> n 1963	<b>െ</b> ന്ന്രിറ്റി 20	07				

# **Unicode Character Properties**

ODE7;SINHALA	DIGIT ON	NE;Nd;0;L;;1;1;1;N;;;;;
ODE8;SINHALA	DIGIT TW	NO;Nd;0;L;;2;2;2;N;;;;;
ODE9;SINHALA	DIGIT TH	HREE;Nd;0;L;;3;3;3;N;;;;;
ODEA; SINHALA	DIGIT FO	<pre>DUR;Nd;0;L;;4;4;4;N;;;;;;</pre>
ODEB;SINHALA	DIGIT FI	IVE;Nd;0;L;;5;5;5;N;;;;;
ODEC;SINHALA	DIGIT SI	IX;Nd;0;L;;6;6;6;N;;;;;
ODED;SINHALA	DIGIT SE	EVEN;Nd;0;L;;7;7;7;N;;;;;
ODEE;SINHALA	DIGIT EI	IGHT;Nd;0;L;;8;8;8;N;;;;;
ODEF;SINHALA	DIGIT NI	INE;Nd;0;L;;9;9;9;N;;;;;
ODF5;SINHALA	NUMBER 1	<pre>FEN;No;0;L;;;;10;N;;;;;</pre>
ODF6;SINHALA	NUMBER 1	<pre>FWENTY;No;0;L;;;;20;N;;;;;</pre>
ODF7;SINHALA	NUMBER 1	<pre>FHIRTY;No;0;L;;;;30;N;;;;;</pre>
ODF8;SINHALA	NUMBER F	FORTY;No;0;L;;;;40;N;;;;;
ODF9;SINHALA	NUMBER F	FIFTY;No;0;L;;;;50;N;;;;;
ODFA;SINHALA	NUMBER S	SIXTY;No;0;L;;;;60;N;;;;;
ODFB;SINHALA	NUMBER S	SEVENTY;No;0;L;;;;70;N;;;;;
ODFC;SINHALA	NUMBER E	EIGHTY;No;0;L;;;;80;N;;;;;
ODFD;SINHALA	NUMBER N	NINETY;No;0;L;;;;90;N;;;;;
ODFE;SINHALA	NUMBER C	<pre>DNE HUNDRED;No;0;L;;;;100;N;;;;;</pre>
ODFF;SINHALA	NUMBER C	<pre>DNE THOUSAND;No;0;L;;;;1000;N;;;;;</pre>

## **Bibliography**

Faulmann, Carl. 1990 (1880). *Das Buch der Schrift*. Frankfurt am Main: Eichborn. ISBN 3-8218-1720-8 Gunasekara, Abraham Mendis. 1986 (1891). *A comprehensive grammar of the Sinhalese language*. New

- Delhi: Asian Educational Services.
- Ifrah, Georges. 1989. *Universalgeschichte der Zahlen*. Translated from the French by Alexander von Platen. Frankfurt: Campus Verlag. ISBN 3-593-34192-1.

Ifrah, Georges. 2000. *The universal history of numbers. Volume 1: The world's first number-systems. Volume 2: The modern number-system.* Translated from the French by David Bellos, E. F. Harding. Sophie Wood, and Ian Monk. London: Harvill Press. ISBN 1-86046-790-3, ISBN 1-86046-791-1

Renou, L., & J. Filliozat. 1953. L'Inde classique: manuel des études indiennes. Hanoi.

	0D8	0D9	0DA	0DB	0DC	0DD	0DE	0DF
0		පා	Ð	ධ	ව	ൂ		
1		Ç	ජ	ත	ୠ	ୁ		
2	ം	ð	ජ		ෂ	ି		ാമ
3	ଃ	වේ	ඣ	Ę	ස	ී		ാ
4		3	kię	ප	හ	୍		Mh
5	<b>a</b>	Q	ଝଟ୍	ඵ	E			で よ
6	ၾာ	ඖ	ජ	ର	3	ि		ß
7	Ŷ		O	හ			ର	S
8	Ê		ඨ	۲		ം	Q	(SA)
9	9		ඩ	ନ		ෙ	M	කි
A	ඊ	ක	ඪ	ය	ଁ	ේ	പ	ଝ୍ଟା
В	Ĉ	6	00	୪		ෙ	M	ର୍ଷଚ୍ୟ
С	Çə	ග	ඬ			ො	Ø	S
D	පෙෘ	な	ろ	C		ෝ	ୟ	S
E	esaa	ඩ	Ċ			ෞ	g	m
F	ප	හ	Ę		ാ	ാ	$\bigcirc$	ଲ୍ୟା

# TABLE XXX - Row 0D: SINHALA

G = 00 P = 00

# TABLE XXX - Row 0D: SINHALA

hex	Name	] [	hex	Name
80 812 33 45 56 78 89 AB & DE F 0 1 2 3 4 5 6 7 8 9 AB C D E F 0 1 2 3 4 5 6 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	IThis position shall not be used) SINHALA SIGN VISARGAYA (This position shall not be used) SINHALA SIGN VISARGAYA SINHALA LETTER AYANNA SINHALA LETTER AYANNA SINHALA LETTER AYANNA SINHALA LETTER AYANNA SINHALA LETTER AYANNA SINHALA LETTER ITANNA SINHALA LETTER IVANNA SINHALA LETTER IVANNA SINHALA LETTER IVANNA SINHALA LETTER IVANNA SINHALA LETTER IVANNA SINHALA LETTER IVANNA SINHALA LETTER IUVANNA SINHALA LETTER IUVANNA SINHALA LETTER IUVANNA SINHALA LETTER IUVANNA SINHALA LETTER IUVANNA SINHALA LETTER IUVYANNA SINHALA LETTER IUVYANNA SINHALA LETTER IUVYANNA SINHALA LETTER IUVYANNA SINHALA LETTER IUVYANNA SINHALA LETTER OYANNA SINHALA LETTER OYANNA SINHALA LETTER OYANNA SINHALA LETTER OYANNA SINHALA LETTER AYANNA SINHALA LETTER OYANNA SINHALA LETTER OYANNA SINHALA LETTER OYANNA SINHALA LETTER OYANNA SINHALA LETTER OYANNA SINHALA LETTER MAHAPAPAANA KAYANNA SINHALA LETTER MAHAPPAANA CAYANNA SINHALA LETTER MAHAPPAANA JAYANNA SINHALA LETTER MAHAPPAANA CAYANNA SINHALA LETTER MAHAPPAANA DAYANNA SINHALA LETTER MAHAPPA		D9 DAB DDC DDE DE E12134556789ABCDEEF01F2F34F5F67F8F9ABCDEFF	SINHALA VOWEL SIGN KOMBUVA SINHALA VOWEL SIGN KOMBUVA HAA AELA-PILLA SINHALA VOWEL SIGN KOMBUVA HAA AELA-PILLA SINHALA VOWEL SIGN KOMBUVA HAA ADGA AELA-PILLA SINHALA VOWEL SIGN KOMBUVA HAA AGAYANUKITTA (This position shall not be used) (This position shall not be used) SINHALA DIGIT TWO SINHALA DIGIT TWO SINHALA DIGIT THREE SINHALA DIGIT TIVE SINHALA DIGIT TIVE SINHALA DIGIT SIX SINHALA DIGIT SIX SINHALA DIGIT SIX SINHALA DIGIT SIX SINHALA OUGH SIX SINHALA VOWEL SIGN DIGA GAETTA-PILLA SINHALA NUMBER TWENTY SINHALA NUMBER TWENTY SINHALA NUMBER THEN SINHALA NUMBER THEN SINHALA NUMBER THEN SINHALA NUMBER THEN SINHALA NUMBER SEVENY SINHALA NUMBER FORTY SINHALA NUMBER SIXTY SINHALA NUMBER SIXTY SINHALA NUMBER ONE THOUSAND

# Figures

Zeichen	Wert	Zeichen	Wert	Zeichen	Wert	Zeichen	Wert	Zeichen	Wert	Zeichen	Wert
		.  <u>i</u>		!				} ;			
¢:	A	පේ	ลเ่	ෂ	gha	ධ	ţha	ತಾ	na	G	la
φD	ā	ම	0	0	ńa	l 🖸 🗄	da	Ö	րո	C	ļa
3	i	5	ō	ତ	tśa	ස	dha	ප	թհո	อ่	va
õ	ī	5.5	£	5	táha	.55	ņa	ଚ	ba	ස	sa
С	u	an a	ē	್	dża	জ	ta	හ	bha	ଓନ	ša
Co	ū	කා	ka	50	džha	Ö	tha	3	ma	3	ša
ల	e	ର	kha	200	ษัล -	Č,	da	ය	ya	හ	ha
હે	ē	හ	¥a	G	ţя	â	dha	6	ra.		

# SINGALESISCH.

Ligaturen.

		1		f		1		ī ·		1	11
25. k		නි	tva	ಕು	11	භූ	bhu	8	มาเ	8	ี่ รî
නු k	ra	Ŝ	$\mathrm{d}\bar{\mathrm{a}}$	ති	ni	ම	ni	မာ	rga	E3	su
<b>25</b> 33 k	ša	8	di	്വ	nu	ම	mu	Ċ	· i	ස්ව	SVa
တီးဖ		ŝ	· du	නි	nva	(3)	mba	A	: Iu	லி	hi
62, ń	ய்க	Ċ,	d <u>e</u>	ප්	- p	25	У	07	lū		
ට ti	i	ŝ	dra	පි	: pi	୍ତ	r	Ğ	v		
ð ti	ī	8	i dri	2	րս	ග	rā	3	vī		
න් t		Ð	ddha	ଚି	b	8	ri	ස්	. 5		

Vokalverbindung: mo kā, mo k

Ceylou war schon den Zeitgenossen Alexanders unter dem Namen T $\alpha \pi \gamma \omega$ - $\beta \Delta \nu \pi$  (nach der Stadt Tambapanni, im Pali Tamraparni) hekannt, Ptolemäus nennt es  $\Sigma \alpha \lambda \omega \pi$  (indisch Sinhala Löwenreich), woraus das arabische Serendib (dib Insel) und das europäische Zeilon oder Ceylon entstand.

Die Schrift schliesst sich der benachbarten *matabarischen* an, nur ist sie zierlicher und sind die Striche mehr geknickt. ihr Grundcharakter ist der der *Pali*, da die buddhistische Religion schon im Jahre 322 v. Chr. eindraug. Eigenthümlich ist die Gestalt des Virama<sup>\*</sup>, welches über die Konsonanten gestellt wird und nicht nur am Ende, sondern auch mitten im Worte steht, wesshalb in dieser Schrift ein Unterstellen der Kousonantenzeichen, wie in der Deranagari- und Palischrift nicht nothwendig ist, und die Ligaturen sieh somit nur auf die Vokalverbindungen beschränken. Dasselbe Zeichen bezeichnet bei e die Länge und macht aus o das aa. Bei b und w ist das Virama durch Verdopplung der Endschleife: B b, D w ersetzt.

and CA SINHALA NUMBER TEN, here mislabelled "0" by Faulmann.



Figure 2. Sample from Gunasekara 1891 showing old Sinhala digits and numbers. Gunasekara notes:

"The Sinhalese had symbols of its own to represent the different numerals, which were in use until the beginning of the present [= 19th] century. Arabic figures are now [= 1891] universally used. For the benefit of the student the old symbols are given in the plate [as shown above]. These are the only figures known in the language for making ordinary calculations, and may be used for expressing any number. But for making astronomical calculations different figures are employed. (Vowel-consonants are even now employed in lieu of figures to indicate the number of leaves in ólá books.)"

No evidence for the astronomical calculation figures has been seen; these could be added to the standard at a later date if required.

	706				APPEN	DYLE 1.					
VALEUR.	GUPTA Valabhī.	KUČĀ, Asie Centrale, Kliotan.	Mss népálais, x1°-x1v° siècle.	Mss jaina.	SINGHALAIS.	VALEUR.	GUPTA Valabhí,	KUČĀ, Asie Centrale, Khotan.	Mss népálais x1°-x1v° siècle.	Mss jaina.	SINGHALAIS.
1	-	* *	2 2 V	م	ଗ	100	स्र	ef.	AA	Eft.	52r
2	=	8 8	G RE (duri)	2	ര	200	4 7	in the	तु आ	展	agy
3	III	999	5 9 (a)	3	Gru	300	x		, जु	灵	argy
4	¥ŧ	华长	的动物	ব্ধা	eg	400	अंभ र	\$		म्	and
5	んち	PA	g a E	Z	Gn	1.000					A
6	<b>45</b>	<b>3</b> H	62	駒	6		TA- MOUL.	MALAY- ĀĻAM.		TA- MOUL.	MALAY. ĀĻAM,
7	15	27	मुभ	IJ	8	1	Б	0	1/80	8×	
8	38	43	\$3	57	R	2	e	2	,1/40	<b>5</b> 4	
9	30	47	83	13	Ì	3	līm	62	3/80	B	
10	44	14.1	rE 37	55	1650	4	P	01	1/20	ч	
11	3.	141 1	र त. ज	M.	ରେମ ହୋ	5	G	3	1/16	128	ഹമ
20	00	8 6	व थ	ส	æS	6	Fre	m	1/10	2	
30	শু থ	C.I	ल ल	ৰ্ন্স	S	7	எ	9	1/8	வற்ற	റഞ
40	Ч	3 3	A SI	প্ন	භිව	8	Э	ч	3/20	F	m
50	C	e c	CC	G	æ	9	<i>6</i> 50	nb	3/16	65L	in
60	31	7	मर्	IJ	BEN	10	Ð	w	1/5	ଇ	
70	2 4	3	ब्रब्	虱	୧୫୧୬୬	100	m	m	1/4	ୟ	6
30	60 00	0	90	0	S	1.000	53	600	1/2	2	ø
00	€9-€9.	G	\$8	8	S	1/320	ണ്ട		3/4	65	ങ

**Figure 3.** Sample from Renou & Filliozat 1953 showing old Sinhala digits and numbers alongside the digits and numbers of other Brahmic scripts.

#### Sinhalese figures

Used mainly in Sri Lanka and in the Maldives as well as in the islands to the north of the latter. (In the north and northwest of Sri Lanka, Tamil figures are also used due to the high number of Tamil people who live in these areas of the island.)

1	2	3	4	5	6	7	8	9	0	Ref.
â	ര	സ	ფ	Gn.	ව	ୠ	ಥ	6		Alwis (de) Charter
ଗ	60	ଚୖୖୖ	ත	ണം	ଭ	3	q	6		Frédéric, DCJ Pihan
୶	ຈ	n	m	m	ଚ	Н	Ţ	୭		Renou and Filliozat
Geo Uso	ograpi ed in S	hical ar ri Lanl	rea (Fi ka, in	g. 24. 2 the Ma	7 and ldives	24. 53 , as we	3): ell as ir	the isla	ands t	to the north of the Maldives.

F1G. 24.22. Current Sinhalese (or Sinhala) numerals

It should be noted that although Sinhalese writing is linked to Dravidian forms of writing (even though it is more stylish, striving as it does towards an ornamental effect), the language of this writing is not Dravidian. Sinhalese is an Indo-European language: "it is a language that belongs to Prakrit (dialects) of 'Middle Indian', as several inscriptions written in *Brâhmî* dating from around the second century BCE show. However, after the fifth century CE, the Sinhalese language, separated

from India's Indo-European languages by the Tamil area, developed in an individual style, as did its writing. The two seem to have changed little since 1250" (L. Frédéric).

There are twenty Sinhalese figures. This number of numerical signs is due to the absence of zero and the fact that the system, which is not based upon the place-value system, uses a specific figure for every ten units, as well as special figures that represent 10, 100 and 1,000 (see Fig. 23.18).

Figure 4. Sample from Ifrah 2000 showing three sources for Sinhala numbers and discussing the history of the figures.

At a later period, the inhabitants of Ceylon went through the same change, but starting from a much better system than those above. They assigned a separate sign not only to every power of 10, but also to each of the nine units and to each of the nine tens, and then applied the same principle as above. In this way, the number 7,659 can be broken down (Fig. 23.18) as

fitst appearar Evpe: B2 (hvb Seed for zero Tapacity for r	ice: 0.600 rid numb sign: No. epicsentat	- 900 CE er-system o Existence o tion: Limite	f the seco f zero sig id (see Cl	onó lype: ar No apter 24	Fig. 23.3 , p.735)	ra Base 10		
_		- Ba	se auml	hers (m	odern so	ipt)		
ଣ	<b>ම</b> ං 2	<b>ଲେ</b> 3	න	G™ ₅	ê ë	ନ୍ଦ୍ର	e s	Ē
ଭ	a	<b>1</b>	<b>6079</b> 0	ଷ୍ଟି	<b>SU</b>	<b>?</b> 33	<b>8</b> 80	<b>6</b>
10			<b>75</b> <b>75</b> 100 (= 10	) <sup>4</sup> )	<b>8</b> 1,000 (-	10-)		
Example: 7.	- — 659	ନ	6	8	ଚ	26	ଛ	୕୕୕ଵୄ
			<b>-</b>	 J(11		100	66	

Her, 23.18. Singhalese number system

Figure 5. Sample from Ifrah 2000 showing the Sinhala digits and numbers.

**Figure 6.** Sample from the 1989 German edition of Ifrah showing the Sinhala digits and numbers, including glyph variants for the tens and for ONE HUNDRED.

#### A. Administrative 1. Title

- Proposal to add archaic numbers for Sinhala to the BMP of the UCS
- 2. Requester's name
- Michael Everson
- 3. Requester type (Member body/Liaison/Individual contribution)
- Individual contribution.
- 4. Submission date
- 2007-01-09
- 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

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.)?

Yes.

2b. If YES, with whom?

Roland Russwurm of ceylon-online.com and sinhala-online.com.

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.

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

**Yes.** 6b. If YES, is a rationale provided?

Yes.

### 6c. If YES, reference

#### Keep with other Sinhala characters.

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?

ANDTEK GmbH Am Soeldnermoos 17 D-85399 Hallbergmoos Germany



Date: 04.01.2007

Your contact: Roland Russwurm **Phone:** +49 (0) 811 9594675 **FAX:** +49 811 9594676 E-Mail: rrusswurm@andtek.com

# **Unicode Standard Draft Proposal**

To whom it may concern,

We represent <u>ceylon-online.com</u> and <u>sinhala-online.com</u>, which offers services for Sri Lankan people and delivers Sinhala language resources for educational and business purposes.

We have reviewed the draft proposal by Michael Everson to encode archaic Sinhala digits and numbers in the Unicode Standard, and would like to endorse it. Until the characters are encoded they can only be represented as graphics.

Kind regards

Roland Russwurm ANDTEK GmbH