

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 of the UCS****Source: Michael Everson****Status: Individual Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Replaces: N1473R****Date: 2007-02-08**

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
0DEB	SINHALA DIGIT FIVE
0DEC	SINHALA DIGIT SIX
0DED	SINHALA DIGIT SEVEN
0DEE	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
0DFE	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:

	෧ 1	෨ 2	෩ 3	෪ 4	෫ 5	෬ 6	෭ 7	෮ 8	෯ 9
෧0	෧෧ 11	෧෨ 12	෧෩ 13	෧෪ 14	෧෫ 15	෧෬ 16	෧෭ 17	෧෮ 18	෧෯ 19
෨0	෨෧ 21	෨෨ 22	෨෩ 23	෨෪ 24	෨෫ 25	෨෬ 26	෨෭ 27	෨෮ 28	෨෯ 29
෩0	෩෧ 31	෩෨ 32	෩෩ 33	෩෪ 34	෩෫ 35	෩෬ 36	෩෭ 37	෩෮ 38	෩෯ 39
෪0	෪෧ 41	෪෨ 42	෪෩ 43	෪෪ 44	෪෫ 45	෪෬ 46	෪෭ 47	෪෮ 48	෪෯ 49
෫0	෫෧ 51	෫෨ 52	෫෩ 53	෫෪ 54	෫෫ 55	෫෬ 56	෫෭ 57	෫෮ 58	෫෯ 59
෬0	෬෧ 61	෬෨ 62	෬෩ 63	෬෪ 64	෬෫ 65	෬෬ 66	෬෭ 67	෬෮ 68	෬෯ 69
෭0	෭෧ 71	෭෨ 72	෭෩ 73	෭෪ 74	෭෫ 75	෭෬ 76	෭෭ 77	෭෮ 78	෭෯ 79
෮0	෮෧ 81	෮෨ 82	෮෩ 83	෮෪ 84	෮෫ 85	෮෬ 86	෮෭ 87	෮෮ 88	෮෯ 89
෯0	෯෧ 91	෯෨ 92	෯෩ 93	෯෪ 94	෯෫ 95	෯෬ 96	෯෭ 97	෯෮ 98	෯෯ 99
෧00	෧෦෧ 101	෧෦෨ 102	෧෦෩ 103	෧෦෪ 104	෧෦෫ 105	෧෦෬ 106	෧෦෭ 107	෧෦෮ 108	෧෦෯ 109
෧෦෧ 111		෧෦෨෦෧ 222		෧෦෩෦෧ 333		෧෦෪෦෧ 444		෧෦෫෦෧ 555	
෧෦෬෦෧ 666		෧෦෭෦෧ 777		෧෦෮෦෧ 888		෧෦෯෦෧ 999			
෧෦෧෦෦ 1000		෧෦෧෦෦෦෧ 1963		෧෦෧෦෦෦෦෦ 2007					

**Unicode Character Properties**

```

ODE7;SINHALA DIGIT ONE;No;0;L;;;1;N;;;;;
ODE8;SINHALA DIGIT TWO;No;0;L;;;2;N;;;;;
ODE9;SINHALA DIGIT THREE;No;0;L;;;3;N;;;;;
ODEA;SINHALA DIGIT FOUR;No;0;L;;;4;N;;;;;
ODEB;SINHALA DIGIT FIVE;No;0;L;;;5;N;;;;;
ODEC;SINHALA DIGIT SIX;No;0;L;;;6;N;;;;;
ODED;SINHALA DIGIT SEVEN;No;0;L;;;7;N;;;;;
ODEE;SINHALA DIGIT EIGHT;No;0;L;;;8;N;;;;;
ODEF;SINHALA DIGIT NINE;No;0;L;;;9;N;;;;;
ODF5;SINHALA NUMBER TEN;No;0;L;;;10;N;;;;;
ODF6;SINHALA NUMBER TWENTY;No;0;L;;;20;N;;;;;
ODF7;SINHALA NUMBER THIRTY;No;0;L;;;30;N;;;;;
ODF8;SINHALA NUMBER FORTY;No;0;L;;;40;N;;;;;
ODF9;SINHALA NUMBER FIFTY;No;0;L;;;50;N;;;;;
ODFA;SINHALA NUMBER SIXTY;No;0;L;;;60;N;;;;;
ODFB;SINHALA NUMBER SEVENTY;No;0;L;;;70;N;;;;;
ODFC;SINHALA NUMBER EIGHTY;No;0;L;;;80;N;;;;;
ODFD;SINHALA NUMBER NINETY;No;0;L;;;90;N;;;;;
ODFE;SINHALA NUMBER ONE HUNDRED;No;0;L;;;100;N;;;;;
ODFF;SINHALA NUMBER ONE THOUSAND;No;0;L;;;1000;N;;;;;
  
```

**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.

TABLE XXX - Row 0D: SINHALA

	0D8	0D9	0DA	0DB	0DC	0DD	0DE	0DF
0		ඌ	ඹ	ඳ	ඹ	ෆ		
1		ඹ	ඳ	න	ශ	ෆ		
2	ං	ඹ	ඳ		ඹ	ි		ාa
3	ං	ඹ	ක	ඳ	ස	ී		ා
4		ඹ	ක	ඹ	හ	ෆ		ෆ
5	ඹ	ඹ	ක	ඹ	ඹ			ඹ
6	ඹ	ඹ	ඳ	ඹ	ඹ	ෆ		ඹ
7	ඹ		ඹ	හ			ඹ	ඹ
8	ඹ		ඹ	ඹ		ා	ඹ	ඹ
9	ඹ		ඹ	ඹ		ි	ඹ	ඹ
A	ඹ	ක	ඹ	ස	්	ේ	ඹ	ඹ
B	ඹ	ඹ	ඹ	ර		ේ	ඹ	ඹ
C	ඹ	ශ	ඹ			ේ	ඹ	ඹ
D	සා	ස	න	ඹ		ේ	ඹ	ඹ
E	සාa	ඹ	ඹ			ේ	ඹ	ඹ
F	ඹ	ශ	ඳ		ා	ා	ඹ	ඹ

G = 00  
P = 00

**TABLE XXX - Row 0D: SINHALA**

hex	Name
80	(This position shall not be used)
81	(This position shall not be used)
82	SINHALA SIGN ANUSVARAYA
83	SINHALA SIGN VISARGAYA
84	(This position shall not be used)
85	SINHALA LETTER AYANNA
86	SINHALA LETTER AAYANNA
87	SINHALA LETTER AEYANNA
88	SINHALA LETTER AEEYANNA
89	SINHALA LETTER IYANNA
8A	SINHALA LETTER IYANNA
8B	SINHALA LETTER UYANNA
8C	SINHALA LETTER UUYANNA
8D	SINHALA LETTER IRUYANNA
8E	SINHALA LETTER IRUYANNA
8F	SINHALA LETTER ILUYANNA
90	SINHALA LETTER ILUUYANNA
91	SINHALA LETTER EYANNA
92	SINHALA LETTER EYANNA
93	SINHALA LETTER AIYANNA
94	SINHALA LETTER OYANNA
95	SINHALA LETTER OYANNA
96	SINHALA LETTER AUANNA
97	(This position shall not be used)
98	(This position shall not be used)
99	(This position shall not be used)
9A	SINHALA LETTER ALPAPRAANA KAYANNA
9B	SINHALA LETTER MAHAAPRAANA KAYANNA
9C	SINHALA LETTER ALPAPRAANA GAYANNA
9D	SINHALA LETTER MAHAAPRAANA GAYANNA
9E	SINHALA LETTER KANTAJA NAASIKYAYA
9F	SINHALA LETTER SANYAKA GAYANNA
A0	SINHALA LETTER ALPAPRAANA CAYANNA
A1	SINHALA LETTER MAHAAPRAANA CAYANNA
A2	SINHALA LETTER ALPAPRAANA JAYANNA
A3	SINHALA LETTER MAHAAPRAANA JAYANNA
A4	SINHALA LETTER TAALUJA NAASIKYAYA
A5	SINHALA LETTER TAALUJA SANYOOGA NAASIKYAYA
A6	SINHALA LETTER SANYAKA JAYANNA
A7	SINHALA LETTER ALPAPRAANA TTAYANNA
A8	SINHALA LETTER MAHAAPRAANA TTAYANNA
A9	SINHALA LETTER ALPAPRAANA DDAYANNA
AA	SINHALA LETTER MAHAAPRAANA DDAYANNA
AB	SINHALA LETTER MUURDHAJA NAYANNA
AC	SINHALA LETTER SANYAKA DDAYANNA
AD	SINHALA LETTER ALPAPRAANA TAYANNA
AE	SINHALA LETTER MAHAAPRAANA TAYANNA
AF	SINHALA LETTER ALPAPRAANA DAYANNA
B0	SINHALA LETTER MAHAAPRAANA DAYANNA
B1	SINHALA LETTER DANTAJA NAYANNA
B2	(This position shall not be used)
B3	SINHALA LETTER SANYAKA DAYANNA
B4	SINHALA LETTER ALPAPRAANA PAYANNA
B5	SINHALA LETTER MAHAAPRAANA PAYANNA
B6	SINHALA LETTER ALPAPRAANA BAYANNA
B7	SINHALA LETTER MAHAAPRAANA BAYANNA
B8	SINHALA LETTER MAYANNA
B9	SINHALA LETTER AMBA BAYANNA
BA	SINHALA LETTER YAYANNA
BB	SINHALA LETTER RAYANNA
BC	(This position shall not be used)
BD	SINHALA LETTER DANTAJA LAYANNA
BE	(This position shall not be used)
BF	(This position shall not be used)
C0	SINHALA LETTER VAYANNA
C1	SINHALA LETTER TAALUJA SAYANNA
C2	SINHALA LETTER MUURDHAJA SAYANNA
C3	SINHALA LETTER DANTAJA SAYANNA
C4	SINHALA LETTER HAYANNA
C5	SINHALA LETTER MUURDHAJA LAYANNA
C6	SINHALA LETTER FAYANNA
C7	(This position shall not be used)
C8	(This position shall not be used)
C9	(This position shall not be used)
CA	SINHALA SIGN AL-LAKUNA
CB	(This position shall not be used)
CC	(This position shall not be used)
CD	(This position shall not be used)
CE	(This position shall not be used)
CF	SINHALA VOWEL SIGN AELA-PILLA
D0	SINHALA VOWEL SIGN KETTI AEDA-PILLA
D1	SINHALA VOWEL SIGN DIGA AEDA-PILLA
D2	SINHALA VOWEL SIGN KETTI IS-PILLA
D3	SINHALA VOWEL SIGN DIGA IS-PILLA
D4	SINHALA VOWEL SIGN KETTI PAA-PILLA
D5	(This position shall not be used)
D6	SINHALA VOWEL SIGN DIGA PAA-PILLA
D7	(This position shall not be used)
D8	SINHALA VOWEL SIGN GAETTA-PILLA

hex	Name
D9	SINHALA VOWEL SIGN KOMBUVA
DA	SINHALA VOWEL SIGN DIGA KOMBUVA
DB	SINHALA VOWEL SIGN KOMBU DEKA
DC	SINHALA VOWEL SIGN KOMBUVA HAA AELA-PILLA
DD	SINHALA VOWEL SIGN KOMBUVA HAA DIGA AELA-PILLA
DE	SINHALA VOWEL SIGN KOMBUVA HAA GAYANUKITTA
DF	SINHALA VOWEL SIGN GAYANUKITTA
E0	(This position shall not be used)
E1	(This position shall not be used)
E2	(This position shall not be used)
E3	(This position shall not be used)
E4	(This position shall not be used)
E5	(This position shall not be used)
E6	(This position shall not be used)
E7	<u>SINHALA DIGIT ONE</u>
E8	<u>SINHALA DIGIT TWO</u>
E9	<u>SINHALA DIGIT THREE</u>
EA	<u>SINHALA DIGIT FOUR</u>
EB	<u>SINHALA DIGIT FIVE</u>
EC	<u>SINHALA DIGIT SIX</u>
ED	<u>SINHALA DIGIT SEVEN</u>
EE	<u>SINHALA DIGIT EIGHT</u>
EF	<u>SINHALA DIGIT NINE</u>
F0	SINHALA VOWEL SIGN DIGA GAETTA-PILLA
F1	SINHALA VOWEL SIGN DIGA GAYANUKITTA
F2	SINHALA PUNCTUATION KUNDDALIYA
F3	(This position shall not be used)
F4	(This position shall not be used)
F5	<u>SINHALA NUMBER TEN</u>
F6	<u>SINHALA NUMBER TWENTY</u>
F7	<u>SINHALA NUMBER THIRTY</u>
F8	<u>SINHALA NUMBER FORTY</u>
F9	<u>SINHALA NUMBER FIFTY</u>
FA	<u>SINHALA NUMBER SIXTY</u>
FB	<u>SINHALA NUMBER SEVENTY</u>
FC	<u>SINHALA NUMBER EIGHTY</u>
FD	<u>SINHALA NUMBER NINETY</u>
FE	<u>SINHALA NUMBER ONE HUNDRED</u>
FF	<u>SINHALA NUMBER ONE THOUSAND</u>

# Figures

## SINGALESISCH.

Zeichen	Wert	Zeichen	Wert	Zeichen	Wert	Zeichen	Wert	Zeichen	Wert	Zeichen	Wert
අ	a	ආ	ai	ඃ	gha	඄	pha	අ	na	ආ	la
ඈ	ā	ඉ	o	ඊ	na	උ	da	ඌ	pa	ඍ	la
ඎ	i	ඏ	ō	ඐ	t̄sa	එ	ḍha	ඒ	pha	ඓ	va
ඔ	ī	ඔ	e	ඒ	t̄sha	ඣ	ṇa	ඥ	ba	ඦ	sa
ඬ	u	ඦ	ē	ඣ	d̄za	ඤ	ta	ඨ	bha	ඩ	śa
ඨ	ū	ට	ka	ඥ	d̄zha	ඩ	tha	ඪ	ma	ණ	śa
ඪ	e	ණ	kha	ඩ	ṇa	ඪ	da	ඬ	ya	ද	ha
ද	ē	ද	ga	ණ	ta	ථ	dha	ද	ra		

### Ligaturen.

ක	k	ත	tva	න	u	භ	bhu	ර	rī	ස	si
ක්‍ර	kra	ද	dā	නි	ni	භි	mī	ර්‍ය	rga	සු	su
ක්‍ෂ	kṣa	දි	dī	නු	nu	භු	mu	ල	l	ස්‍ර	sva
ග	g	දු	du	භ	nva	භ්‍ර	mba	ලු	lu	හි	hi
ග්‍ර	grā	ද්‍ර	dṛ	භ	p	ර්‍ය	y	ලූ	lū		
භ	ḥi	ද්‍රා	dra	භ	pi	ර්‍ය	r	ව	v		
භි	ḥī	ද්‍රි	dri	භ	pu	ර්‍ය	rā	වි	vī		
භ්‍ර	ḥr	ද්‍ර්‍ය	ḍdha	භ	b	ර්‍ය	ri	ස්	s		

Vokalverbindung: කා kā, කි ki, කී kī, කු ku, කූ kū, කේ ke, කේ kē, කො kai, කො ko, කෝ kō, කෙ ke, කෙ kē, ක්‍ර kṛ, ක්‍ර kṛ, ක්‍ර kṛ, ක්‍ර kṛ, ක්‍ර kṛ, ක්‍ර kṛ, ක්‍ර kṛ, ක්‍ර kṛ, ක්‍ර kṛ, ක්‍ර kṛ.

Ziffern: 1, 2, 3, 4, 5, 6, 7, 8, 9, 0.

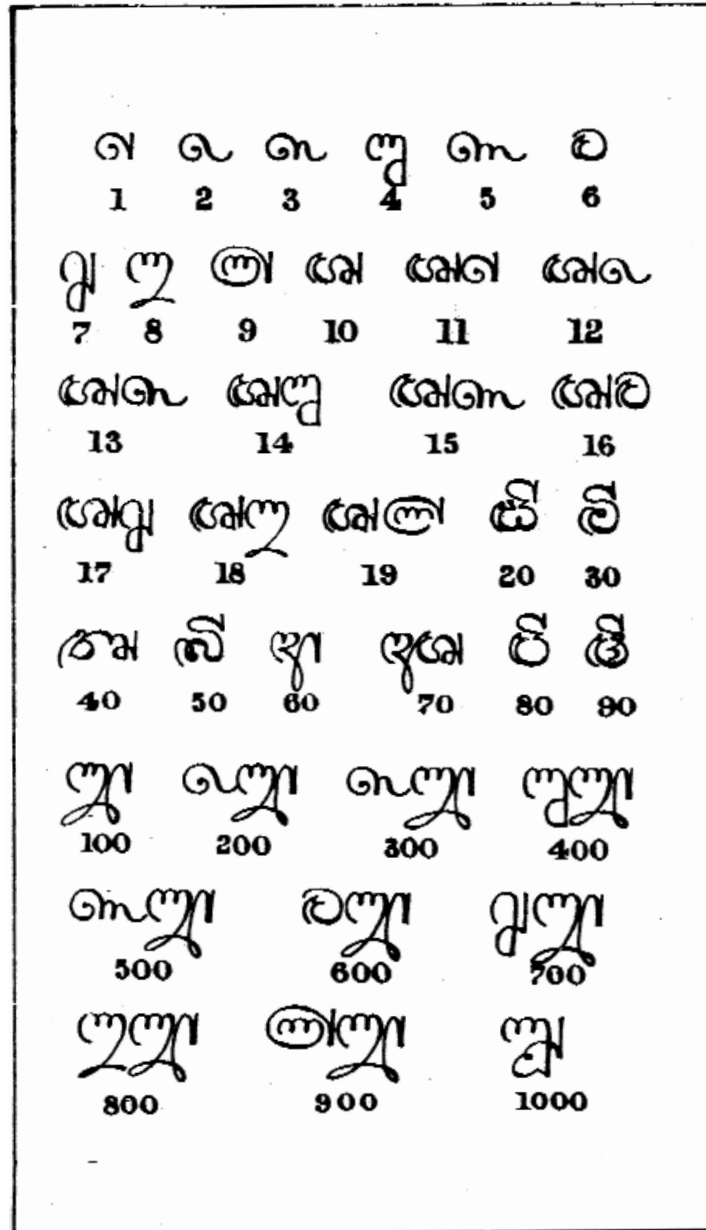
Ceylon war schon den Zeitgenossen Alexanders unter dem Namen *Taxos-βέρος* (nach der Stadt *Tambapanni*, im Pali *Tamraparni*) bekannt, Ptolemäus nennt es *Σαλιζή* (indisch *Sinhala* Löwenreich), woraus das arabische *Serendib* (*dib* Insel) und das europäische *Zeilon* oder *Ceylon* entstand.

Die Schrift schliesst sich der benachbarten *malabarischen* 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. eindrang.

Eigenthümlich ist die Gestalt des *Virama*, welches über die Konsonanten gestellt wird und nicht nur am Ende, sondern auch mitten im Worte steht, wesshalb in dieser Schrift ein Unterstellen der Konsonantenzeichen, wie in der *Devanagari*- und *Palis*chrift nicht nothwendig ist, und die Ligaturen sich 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: *බ්* *b*, *ව්* *w* ersetzt.

**Figure 1.** Sample from Faulmann 1880 showing ෧ SINHALA DIGIT ONE, ෨ SINHALA DIGIT TWO, ෩ SINHALA DIGIT THREE, ෪ SINHALA DIGIT FOUR, ෫ SINHALA DIGIT FIVE, ෬ SINHALA DIGIT SIX, ෭ SINHALA DIGIT SEVEN, ෮ SINHALA DIGIT EIGHT, ෯ SINHALA DIGIT NINE, and ෧෦ SINHALA NUMBER TEN, here mislabelled “0” by Faulmann.

*Plate III.*



*Lith. Surveyor General's Office, Colombo. N° 226.*

**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.

VALEUR.	GUPTA Valabhi.	KUCĀ, Asie Centrale, Khotan.	Miss népalais, x <sup>e</sup> -xiv <sup>e</sup> siècle.	Miss jaina.	SINGHALAIS.	VALEUR.	GUPTA Valabhi.	KUCĀ, Asie Centrale, Khotan.	Miss népalais, x <sup>e</sup> -xiv <sup>e</sup> siècle.	Miss jaina.	SINGHALAIS.
1	𑀓	𑀓	𑀓	𑀓	𑀓	100	𑀓	𑀓	𑀓	𑀓	𑀓
2	𑀔	𑀔	𑀔	𑀔	𑀔	200	𑀔	𑀔	𑀔	𑀔	𑀔
3	𑀕	𑀕	𑀕	𑀕	𑀕	300	𑀕	𑀕	𑀕	𑀕	𑀕
4	𑀖	𑀖	𑀖	𑀖	𑀖	400	𑀖	𑀖	𑀖	𑀖	𑀖
5	𑀗	𑀗	𑀗	𑀗	𑀗	1.000	𑀗	𑀗	𑀗	𑀗	𑀗
6	𑀘	𑀘	𑀘	𑀘	𑀘						
7	𑀙	𑀙	𑀙	𑀙	𑀙	1	𑀙	𑀙	1/80	𑀙	𑀙
8	𑀚	𑀚	𑀚	𑀚	𑀚	2	𑀚	𑀚	1/40	𑀚	𑀚
9	𑀛	𑀛	𑀛	𑀛	𑀛	3	𑀛	𑀛	3/80	𑀛	𑀛
10	𑀜	𑀜	𑀜	𑀜	𑀜	4	𑀜	𑀜	1/20	𑀜	𑀜
11	𑀝	𑀝	𑀝	𑀝	𑀝	5	𑀝	𑀝	1/16	𑀝	𑀝
20	𑀞	𑀞	𑀞	𑀞	𑀞	6	𑀞	𑀞	1/10	𑀞	𑀞
30	𑀟	𑀟	𑀟	𑀟	𑀟	7	𑀟	𑀟	1/8	𑀟	𑀟
40	𑀠	𑀠	𑀠	𑀠	𑀠	8	𑀠	𑀠	3/20	𑀠	𑀠
50	𑀡	𑀡	𑀡	𑀡	𑀡	9	𑀡	𑀡	3/16	𑀡	𑀡
60	𑀢	𑀢	𑀢	𑀢	𑀢	10	𑀢	𑀢	1/5	𑀢	𑀢
70	𑀣	𑀣	𑀣	𑀣	𑀣	100	𑀣	𑀣	1/4	𑀣	𑀣
80	𑀤	𑀤	𑀤	𑀤	𑀤	1.000	𑀤	𑀤	1/2	𑀤	𑀤
90	𑀥	𑀥	𑀥	𑀥	𑀥	1/320 1/160	𑀥	𑀥	3/4	𑀥	𑀥

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.
ඈ	ල	෧	ඔ	෧	ච	ඔ	ඔ	ඔ	ඔ	Alwis (de) Charter Frédéric, DCJ Pihan Renou and Filliozat
෧	෧	෧	ඔ	෧	ච	ඔ	ඔ	ඔ	ඔ	
෧	෧	෧	ඔ	෧	ච	ඔ	ඔ	ඔ	ඔ	
Geographical area (Fig. 24. 27 and 24. 53): Used in Sri Lanka, in the Maldives, as well as in the islands to the north of the Maldives.										

FIG. 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 *Brâhmi* 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" (I. 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

$$7 \times 1,000 + 6 \times 100 + 50 + 9.$$

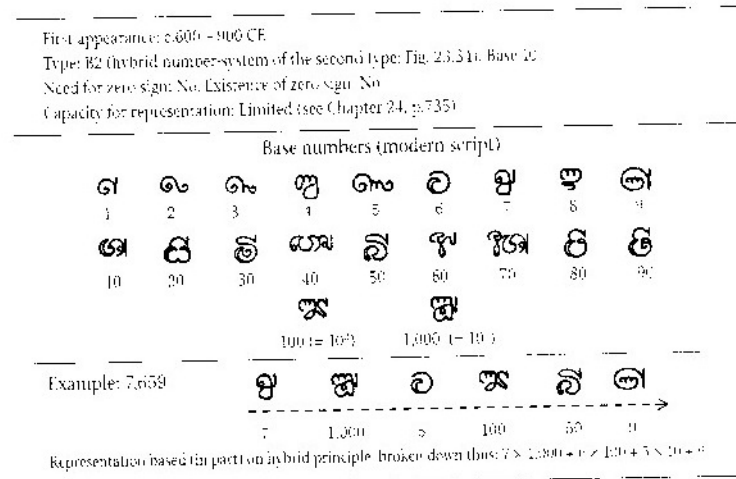


Fig. 23.18. Sinhalese number system

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

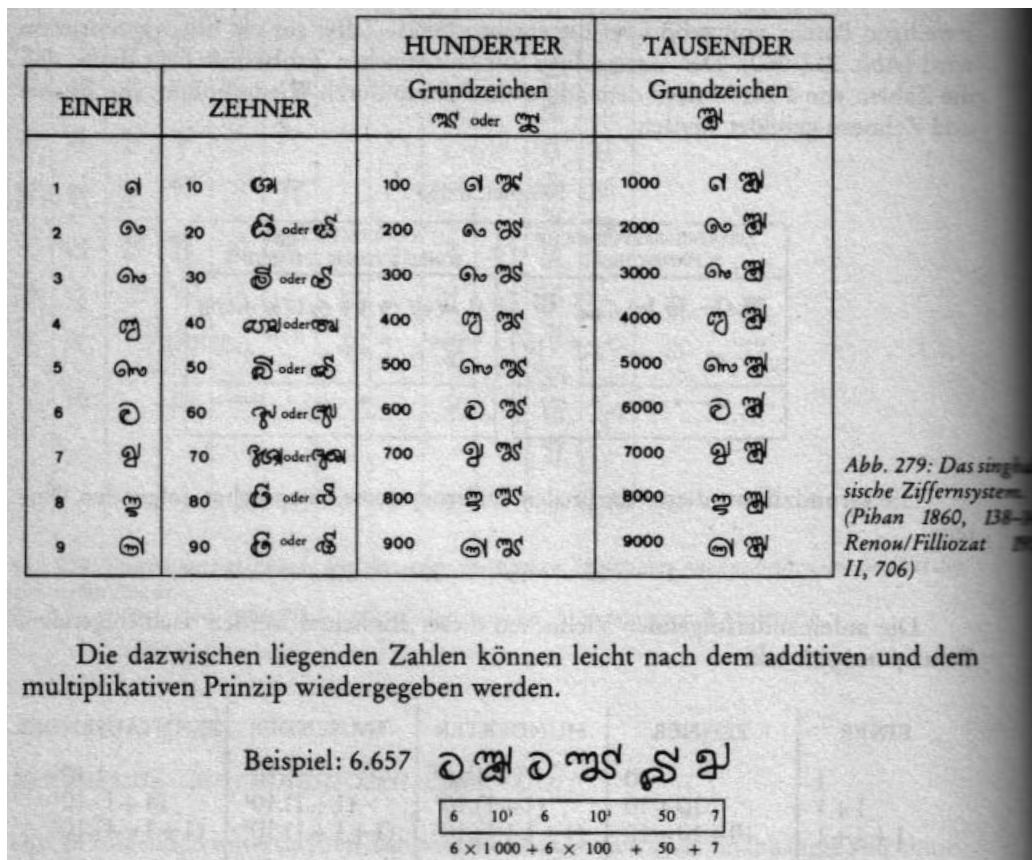


Figure 6. Sample from the 1989 German edition of Ifrah showing the Sinhala digits and numbers, including glyph variants for the tens and for ONE HUNDRED.

## **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

**2007-02-08**

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**

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

**A letter of support is appended to the end of this document.**

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?