

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION  
ORGANISATION INTERNATIONALE DE NORMALISATION  
ISO/IEC JTC 1/SC 2/WG 2**

**Universal Multiple-Octet Coded Character Set  
(UCS)**

**ISO/IEC JTC 1/SC 2 *N*\_\_\_\_\_**  
**ISO/IEC JTC 1/SC 2/WG 2 *N3888***  
**2010-09-1**

|                      |  |
|----------------------|--|
| <b>Doc Type:</b>     | <b>Working Group Document</b>  |
| <b>Title:</b>        | <b>Proposal to include Sinhala Numerals to the BMP and SMP of the UCS</b>    |
| <b>Source:</b>       | <b>Sri Lanka Standards Institution (SLSI)</b>                                |
| <b>Status:</b>       | <b>National Body Contribution</b>  |
| <b>Reference</b>     | <b>See the reference section in the document</b>                             |
| <b>Replaces:</b>     | <b>N3195R</b>  |
| <b>Action:</b>       | <b>For consideration by ISO/IEC JTC 1/SC 2/WG 2 and UTC</b>                  |
| <b>Distribution:</b> | <b>ISO/IEC JTC 1/SC 2/WG 2, ISO/IEC JTC 1/SC 2 and Liaison Organizations</b> |

**1.** Introduction

This paper presents a proposal for encoding Sinhala numerals in Universal Character Set (UCS). Since British occupation of Kandy in 1815, Sinhala scribes had dropped the common use of numerals in Sinhala. Although a few scholars had recorded the existence of Sinhala numerals after 1815, a comprehensive research was required to establish the past existence and precise shapes of these numerals prior to encoding.

The proposal, L2/07-002R (ISO/IEC JTC1/SC2/WG2 N3195R), which was submitted by Mr. Michael Everson to the Unicode consortium for encoding a set of numerals which he claimed were Sinhala numerals, initiated research into Sinhala numerals and numerations. Through the research which was carried out by Mr. Harsha Wijayawardhana of University of Colombo School of Computing (UCSC) under the aegis of the Information and Communication Technology Agency (ICTA) of Sri Lanka, it was discovered that other than the set of numerals submitted by Mr. Michael Everson for encoding, there were four other sets which were commonly used by Sinhala scribes, namely Sinhala Lith Illakkam (Astrological Digits) mainly used for writing horoscopes; Swara, (Sinhala Consonant and vowel modifier based numerals); Katapayadia, a special Sinhala character based numeral set which was used for inscribing years in astrological writing, in ancient Ola and in rock inscriptions; and word based Booththa Anka or Samkaya used in Sinhala Poetry. Mr. Wijayawardhana identified the numeral set which was submitted by Mr. Michael Everson to UCS as Sinhala Illakkam (Sinhala Archaic Numbers).

Subsequently, Prof. K.D. Paranvithana of Raja Rata University, Sri Lanka, and Mr. Harsha Wijayawardhana carried out further research and the findings were presented at the National Archaeological Symposium held in July 2009, in Colombo, Sri Lanka, organized by the Department of Archaeology. The Synopsis of the paper was published in the Volume II of the Symposium's proceedings. In October 2009, Mr. Harsha Wijayawardhana authored a book titled "Numerations in the Sinhala Language".

## 2. Brief overview of the Sinhala Script

The Sinhala Script had evolved from the Southern Brahmi Script from which other scripts such as Telugu and Oriya of South India had evolved. Sinhala belongs to the Indo European family of languages and has a close kinship with northern Indian Languages such as Hindi, Bengali, and Gujarati etc. The Sinhala script was subsequently influenced by Grantha writing of Southern India and shapes of Sinhala letters had remained somewhat static with a few changes since 1250 AD. Sinhala Script has several unique characteristics such as two additional vowels found after the first two vowels, Ae and long Ae (Ae:) which allow words such as “apple” to be written in Sinhala easily. These two vowels are not found in other Indic languages. Sinhala script sans numerals was encoded in UCS in 1997.

## 3. Outcome of the Research

It was found in the research into Sinhala numerals that prior to the occupation of the Kandyan kingdom by the British in 1815, Sinhala scribes had used five sets of numerals:

- i. Sinhala Illakkam: This numeral set does not have a zero and had been commonly used to record numbers especially in the Royal Court. **Since this set of numerals was not used after 1815, these numerals are classified under the category Historical digits and numbers in the encoding.** It has separate symbols for 10,20,30,40,50,60,70,80,100, 1000.
- ii. Sinhala Lith Illakkam: This numeral set was widely used for calculations especially in Astrology and was continuously used well into the 20<sup>th</sup> century, mostly for writing positions of planets in horoscopes. There is some evidence that this set was used for writing page numbers and it may have been commonly used for giving numbers in inscriptions. and Identification of these require more research. Lith Illakkam has a zero and zero place holder concept. For zero, Sinhala Halant or Virama is used in this set.
- iii. Katapayadia: This is a special set of numerals which use Sinhala letters. Each number has several Sinhala letters and a word is made out of letters for a number and a number is depicted from right to left whereas the word can be read from left to right. Some of the Katapyadia based numbers are found in old Tamil rock inscriptions located in Pollonnaruwa. These inscriptions had been created at the time of occupation of Chola. Katapyadia based numbers become more prevalent in Sinhala inscriptions after Pollonnaruwa, Dambedeniya and Kotte periods when Southern Indian influence on Sri Lanka reached its zenith.
- iv. Swara: These are Sinhala consonants with vowel modifiers. These were used primarily for page numbering of Ola but there is evidence that these were used for general numbering of items but not in calculations.
- v. Bootha Anka or Samkaya: Numbers are depicted by certain Sinhala nouns and a method of depicting numbers by words was used primarily by poets. This numbering system had come into Sinhala from Sanskrit and the same objects which were used for numbers in Sanskrit were used in Sinhala as well by employing equivalent Sinhala nouns.

#### 4. Conclusion and follow up by the Local Language Working Group (LLWG) of ICTA

ICTA's Local Language Working Group, after comprehensively reviewing the research and after further extensive discussions and consultations with other experts and stakeholders agreed that:

- i. Sinhala Lith Illakkam (Sinhala Astrological numbers), digits 0-9 to be included in the Basic Multilingual Plane (Sinhala block 0Dxx) of the Unicode standard from 0DE6- 0DFF.
- ii. Sinhala Illakkam or Sinhala Archaic numbers (digits: 1-9, numbers:10-90, 100 and 1000) be included in the supplementary area of the Unicode standard and UTC and recommended that these numerals to be encoded from U+111E0 - U+111FF.

These recommendations were made by LLWG taking into consideration that the set Sinhala Lith Illakkam was continuously used until the early part of 20<sup>th</sup> century and Sinhala Archaic Numerals ceased to be used after 1815.

Note: - Sri Lanka submitted a preliminary proposal to UTC, and UTC agreed that Sinhala Illakkam or Sinhala Archaic Numbers to be encoded from U+111E0 – U+111FF. UTC accepted in principle as recommended by the LLWG of ICTA that Lith Illakkam (Sinhala Astrological Digits) to be placed in the BMP area starting from 0DE6).

The proposals for encoding were presented to the Sectoral Committee on IT, of the Sri Lanka Standards Institution (SLSI). This Committee recommends Sri Lanka Standards to SLSI's Council (Board). The Committee approved in principle that the numerals be added to the Sinhala character encoding given in the Sri Lanka Standards as Sinhala Character Code for Information Interchange, SLS 1134 : 2004.

#### 5. Shapes of Sinhala Illakkam (Sinhala Archaic numbers and Sinhala Lith Illakkam (Astrological numbers))

- i. Sinhala Illakkam (Sinhala Archaic Numbers)

Shapes of Sinhala Illakkam were given by Mendis Gunasekera in his famous book 'A comprehensive Grammar of Sinhala Language'. And also the shapes of Sinhala Illakkam were depicted in W.A De Silva's 'Catalogue of Palm leaf manuscripts in the library of Colombo Museum'. Subsequently, the author discovered Sinhala Illakkam in the Kandyan Convention which was signed between the British and Kandyan Chieftains in 1815. The following Sinhala Illakkam were taken from the Kandyan convention:

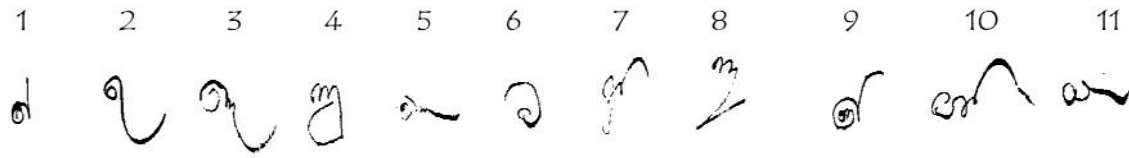


Figure 1 Sinhala Illakkam in the Kandyan Convention

Mr. Harsha Wijayawardhana determined shapes of Sinhala Illakkam based upon all the above sources the shapes and various combination of Sinhala Illakkam are given below:

|          |          |          |           |           |           |
|----------|----------|----------|-----------|-----------|-----------|
|          | ඞ<br>1   | ඞ<br>2   | ඞ<br>3    | ඞ<br>4    | ඞ<br>5    |
| ඞ<br>6   | ඞ<br>7   | ඞ<br>8   | ඞ<br>9    | ඞ<br>10   | ඞ<br>11   |
| ඞ<br>12  | ඞ<br>13  | ඞ<br>14  | ඞ<br>15   | ඞ<br>16   | ඞ<br>17   |
| ඞ<br>18  | ඞ<br>19  | ඞ<br>20  | ඞ<br>21   | ඞ<br>22   | ඞ<br>23   |
| ඞ<br>24  | ඞ<br>25  | ඞ<br>26  | ඞ<br>27   | ඞ<br>28   | ඞ<br>29   |
| ඞ<br>30  | ඞ<br>31  | ඞ<br>32  | ඞ<br>33   | ඞ<br>34   | ඞ<br>35   |
| ඞ<br>36  | ඞ<br>37  | ඞ<br>38  | ඞ<br>39   | ඞ<br>40   | ඞ<br>41   |
| ඞ<br>42  | ඞ<br>43  | ඞ<br>44  | ඞ<br>45   | ඞ<br>46   | ඞ<br>47   |
| ඞ<br>48  | ඞ<br>49  | ඞ<br>50  | ඞ<br>60   | ඞ<br>70   | ඞ<br>80   |
| ඞ<br>90  | ඞ<br>100 | ඞ<br>101 | ඞ<br>102  | ඞ<br>103  | ඞ<br>104  |
| ඞ<br>105 | ඞ<br>106 | ඞ<br>107 | ඞ<br>108  | ඞ<br>109  | ඞ<br>110  |
| ඞ<br>111 | ඞ<br>222 | ඞ<br>333 | ඞ<br>444  | ඞ<br>555  | ඞ<br>666  |
| ඞ<br>777 | ඞ<br>888 | ඞ<br>999 | ඞ<br>1000 | ඞ<br>1987 | ඞ<br>2010 |

ii. Sinhala Lith Illakkam

Three versions of Sinhala Lith Illakkam were discovered and some form of the Halant had been used to depict the zero in all three versions. Mendis Gunasekera had stated in one of his papers that the symbol for zero in Arabic numerals can be used in place of the Sinhala Virama or halant in Sinhala Lith Illakkam.

|   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|
| ආ | ඈ | ඉ | ඊ | උ | ඌ | ඍ | ඎ | ඏ | ඐ |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

|      |      |      |      |      |      |
|------|------|------|------|------|------|
| ආ    | ඈ    | ඉ    | ඊ    | උ    | ඌ    |
| 0    | 1    | 2    | 3    | 4    | 5    |
| ඍ    | ඎ    | ඏ    | ඐ    | එ    | ඒ    |
| 6    | 7    | 8    | 9    | 10   | 11   |
| උආ   | උඈ   | උඉ   | උඊ   | උඋ   | උඌ   |
| 12   | 13   | 14   | 15   | 16   | 17   |
| උඍ   | උඎ   | උඏ   | උඐ   | උඑ   | උඒ   |
| 18   | 19   | 20   | 21   | 22   | 23   |
| උඓ   | උඔ   | උඕ   | උඖ   | උ඗   | උ඘   |
| 24   | 25   | 26   | 27   | 28   | 29   |
| උ඙   | උක   | උඛ   | උඛආ  | උඛඈ  | උඛඉ  |
| 30   | 31   | 32   | 33   | 34   | 35   |
| උඛඊ  | උඛඋ  | උඛඌ  | උඛඍ  | උඛඎ  | උඛඏ  |
| 36   | 37   | 38   | 39   | 40   | 41   |
| උඛඐ  | උඛඑ  | උඛඒ  | උඛඓ  | උඛඔ  | උඛඕ  |
| 42   | 43   | 44   | 45   | 46   | 47   |
| උඛඖ  | උඛ඗  | උඛ඘  | උඛ඙  | උඛක  | උඛඛ  |
| 48   | 49   | 50   | 60   | 70   | 80   |
| උඛඞ  | උඛඟ  | උඛච  | උඛඡ  | උඛජ  | උඛඣ  |
| 90   | 100  | 101  | 102  | 103  | 104  |
| උඛඤ  | උඛඦ  | උඛට  | උඛඨ  | උඛඩ  | උඛඪ  |
| 105  | 106  | 107  | 108  | 109  | 110  |
| උඛණ  | උඛඬ  | උඛත  | උඛඣආ | උඛඣඈ | උඛඣඉ |
| 111  | 222  | 333  | 444  | 555  | 666  |
| උඛඣඊ | උඛඣඋ | උඛඣඌ | උඛඣඍ | උඛඣඎ | උඛඣඏ |
| 777  | 888  | 999  | 1000 | 1987 | 2010 |

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

**Various signs**

- 0D82 𑀮 SINHALA SIGN ANUSVARAYA  
= anusvara
- 0D83 𑀭 SINHALA SIGN VISARGAYA  
= visarga

**Independent vowels**

- 0D85 𑀡 SINHALA LETTER AYANNA  
= sinhala letter a
- 0D86 𑀢 SINHALA LETTER AAYANNA  
= sinhala letter aa
- 0D87 𑀣 SINHALA LETTER AHEYANNA  
= sinhala letter ae
- 0D88 𑀤 SINHALA LETTER AEEYANNA  
= sinhala letter aae
- 0D89 𑀥 SINHALA LETTER IYANNA  
= sinhala letter i
- 0D8A 𑀦 SINHALA LETTER IYANNA  
= sinhala letter ii
- 0D8B 𑀧 SINHALA LETTER UYANNA  
= sinhala letter u
- 0D8C 𑀨 SINHALA LETTER UUYANNA  
= sinhala letter uu
- 0D8D 𑀩 SINHALA LETTER IRUYANNA  
= sinhala letter vocalic r
- 0D8E 𑀪 SINHALA LETTER IRUYANNA  
= sinhala letter vocalic rr
- 0D8F 𑀫 SINHALA LETTER ILUYANNA  
= sinhala letter vocalic l
- 0D90 𑀬 SINHALA LETTER ILUYANNA  
= sinhala letter vocalic ll
- 0D91 𑀭 SINHALA LETTER EYANNA  
= sinhala letter e
- 0D92 𑀮 SINHALA LETTER EYANNA  
= sinhala letter ee
- 0D93 𑀯 SINHALA LETTER AIYANNA  
= sinhala letter ai
- 0D94 𑀰 SINHALA LETTER OYANNA  
= sinhala letter o
- 0D95 𑀱 SINHALA LETTER OYANNA  
= sinhala letter oo
- 0D96 𑀲 SINHALA LETTER AUYANNA  
= sinhala letter au

**Consonants**

- 0D9A 𑀳 SINHALA LETTER ALPAPRAANA KAYANNA  
= sinhala letter ka
- 0D9B 𑀴 SINHALA LETTER MAHAAPRAANA KAYANNA  
= sinhala letter kha
- 0D9C 𑀵 SINHALA LETTER ALPAPRAANA GAYANNA  
= sinhala letter ga
- 0D9D 𑀶 SINHALA LETTER MAHAAPRAANA GAYANNA  
= sinhala letter gha
- 0D9E 𑀷 SINHALA LETTER KANTAJA NAASIKYAYA  
= sinhala letter nga
- 0D9F 𑀸 SINHALA LETTER SANYAKA GAYANNA  
= sinhala letter nnga
- 0DA0 𑀹 SINHALA LETTER ALPAPRAANA CAYANNA  
= sinhala letter ca
- 0DA1 𑀺 SINHALA LETTER MAHAAPRAANA CAYANNA  
= sinhala letter cha
- 0DA2 𑀻 SINHALA LETTER ALPAPRAANA JAYANNA  
= sinhala letter ja

- 0DA3 𑀼 SINHALA LETTER MAHAAPRAANA JAYANNA  
= sinhala letter jha
- 0DA4 𑀽 SINHALA LETTER TAALUJA NAASIKYAYA  
= sinhala letter nya
- 0DA5 𑀾 SINHALA LETTER TAALUJA SANYOOGA  
NAASIKYAYA  
= sinhala letter jnya
- 0DA6 𑀿 SINHALA LETTER SANYAKA JAYANNA  
= sinhala letter nyja
- 0DA7 𑁀 SINHALA LETTER ALPAPRAANA TTAYANNA  
= sinhala letter tta
- 0DA8 𑁁 SINHALA LETTER MAHAAPRAANA TTAYANNA  
= sinhala letter ttha
- 0DA9 𑁂 SINHALA LETTER ALPAPRAANA DDAYANNA  
= sinhala letter dda
- 0DAA 𑁃 SINHALA LETTER MAHAAPRAANA DDAYANNA  
= sinhala letter ddha
- ODAB 𑁄 SINHALA LETTER MUURDHAJA NAYANNA  
= sinhala letter nna
- ODAC 𑁅 SINHALA LETTER SANYAKA DDAYANNA  
= sinhala letter ndda
- ODAD 𑁆 SINHALA LETTER ALPAPRAANA TAYANNA  
= sinhala letter ta
- ODAE 𑁇 SINHALA LETTER MAHAAPRAANA TAYANNA  
= sinhala letter tha
- ODAF 𑁈 SINHALA LETTER ALPAPRAANA DAYANNA  
= sinhala letter da
- 0DB0 𑁉 SINHALA LETTER MAHAAPRAANA DAYANNA  
= sinhala letter dha
- 0DB1 𑁊 SINHALA LETTER DANTAJA NAYANNA  
= sinhala letter na
- 0DB2 𑁋 <reserved>
- 0DB3 𑁌 SINHALA LETTER SANYAKA DAYANNA  
= sinhala letter nda
- 0DB4 𑁍 SINHALA LETTER ALPAPRAANA PAYANNA  
= sinhala letter pa
- 0DB5 𑁎 SINHALA LETTER MAHAAPRAANA PAYANNA  
= sinhala letter pha
- 0DB6 𑁏 SINHALA LETTER ALPAPRAANA BAYANNA  
= sinhala letter ba
- 0DB7 𑁐 SINHALA LETTER MAHAAPRAANA BAYANNA  
= sinhala letter bha
- 0DB8 𑁑 SINHALA LETTER MAYANNA  
= sinhala letter ma
- 0DB9 𑁒 SINHALA LETTER AMBA BAYANNA  
= sinhala letter mba
- 0DBA 𑁓 SINHALA LETTER YAYANNA  
= sinhala letter ya
- 0DBB 𑁔 SINHALA LETTER RAYANNA  
= sinhala letter ra
- 0DBC 𑁕 <reserved>
- 0DBD 𑁖 SINHALA LETTER DANTAJA LAYANNA  
= sinhala letter la  
• dental
- 0DBE 𑁗 <reserved>
- 0DBF 𑁘 <reserved>
- 0DC0 𑁙 SINHALA LETTER VAYANNA  
= sinhala letter va
- 0DC1 𑁚 SINHALA LETTER TAALUJA SAYANNA  
= sinhala letter sha
- 0DC2 𑁛 SINHALA LETTER MUURDHAJA SAYANNA  
= sinhala letter ssa  
• retroflex

- ODC3 ඩ SINHALA LETTER DANTAJA SAYANNA  
= sinhala letter sa  
• dental
- ODC4 ඩ SINHALA LETTER HAYANNA  
= sinhala letter ha
- ODC5 ඩ SINHALA LETTER MUURDHAJA LAYANNA  
= sinhala letter lla  
• retroflex
- ODC6 ඩ SINHALA LETTER FAYANNA  
= sinhala letter fa

## Sign

- ODCA ඩ SINHALA SIGN AL-LAKUNA  
= virama

## Dependent vowel signs

- ODCF ඩ SINHALA VOWEL SIGN AELA-PILLA  
= sinhala vowel sign aa
- ODD0 ඩ SINHALA VOWEL SIGN KETTI AEDA-PILLA  
= sinhala vowel sign ae
- ODD1 ඩ SINHALA VOWEL SIGN DIGA AEDA-PILLA  
= sinhala vowel sign aae
- ODD2 ඩ SINHALA VOWEL SIGN KETTI IS-PILLA  
= sinhala vowel sign i
- ODD3 ඩ SINHALA VOWEL SIGN DIGA IS-PILLA  
= sinhala vowel sign ii
- ODD4 ඩ SINHALA VOWEL SIGN KETTI PAA-PILLA  
= sinhala vowel sign u
- ODD5 ඩ <reserved>
- ODD6 ඩ SINHALA VOWEL SIGN DIGA PAA-PILLA  
= sinhala vowel sign uu
- ODD7 ඩ <reserved>
- ODD8 ඩ SINHALA VOWEL SIGN GAETTA-PILLA  
= sinhala vowel sign vocalic r
- ODD9 ඩ SINHALA VOWEL SIGN KOMBUVA  
= sinhala vowel sign e
- ODDA ඩ SINHALA VOWEL SIGN DIGA KOMBUVA  
= sinhala vowel sign ee  
≡ ODD9 ඩ ODC A ඩ
- ODDB ඩ 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.*

- ODDC ඩ SINHALA VOWEL SIGN KOMBUVA HAA AELA-PILLA  
= sinhala vowel sign o  
≡ ODD9 ඩ ODCF ඩ
- ODDD ඩ SINHALA VOWEL SIGN KOMBUVA HAA DIGA AELA-PILLA  
= sinhala vowel sign oo  
≡ ODDC ඩ ODC A ඩ
- ODDE ඩ SINHALA VOWEL SIGN KOMBUVA HAA GAYANUKITTA  
= sinhala vowel sign au  
≡ ODD9 ඩ ODDF ඩ

## Dependent vowel sign

- ODDF ඩ SINHALA VOWEL SIGN GAYANUKITTA  
= sinhala vowel sign vocalic l

## Sinhala Lith Illakkam or Digits (Sinhala Astrological Digits)

*These digits were primarily used for writing horoscopes after the fall of the Kandyan Kingdom though there is evidence that they were used for other purposes such as writing page numbers etc. The tradition of writing degrees and minutes of zodiac signs in horoscopes continued into the 20th century using different versions of Lith Digits. Prior to the fall of Kandyan Kingdom all calculations were carried out using Lith digits.*

- ODE6 ඩ SINHALA LITH DIGIT ZERO  
= sinhala astrological digit zero
- ODE7 ඩ SINHALA LITH DIGIT ONE  
= sinhala astrological digit one
- ODE8 ඩ SINHALA LITH DIGIT TWO  
= sinhala astrological digit two
- ODE9 ඩ SINHALA LITH DIGIT THREE  
= sinhala astrological digit three
- ODEA ඩ SINHALA LITH DIGIT FOUR  
= sinhala astrological digit four
- ODEB ඩ SINHALA LITH DIGIT FIVE  
= sinhala astrological digit five
- ODEC ඩ SINHALA LITH DIGIT SIX  
= sinhala astrological digit six
- ODED ඩ SINHALA LITH DIGIT SEVEN  
= sinhala astrological digit seven
- ODEE ඩ SINHALA LITH DIGIT EIGHT  
= sinhala astrological digit eight
- ODEF ඩ SINHALA LITH DIGIT NINE  
= sinhala astrological digit nine


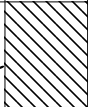
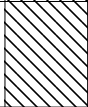
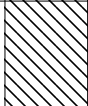
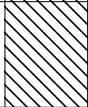
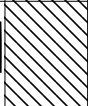
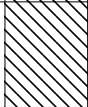
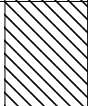
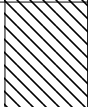
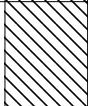
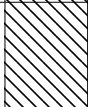
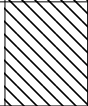
## Additional dependent vowel signs

- ODF2 ඩ SINHALA VOWEL SIGN DIGA GAETTA-PILLA  
= sinhala vowel sign vocalic rr
- ODF3 ඩ SINHALA VOWEL SIGN DIGA GAYANUKITTA  
= sinhala vowel sign vocalic ll

## Punctuation

- ODF4 ඩ SINHALA PUNCTUATION KUNDDALIYA



|   | 111E  | 111F  |
|---|---|---|
| 0 |  | අර්ථ<br>111F0   |
| 1 | ෧<br>111E1  | ඒ<br>111F1  |
| 2 | ෨<br>111E2  | ඔ<br>111F2  |
| 3 | ෩<br>111E3  | ඥ<br>111F3  |
| 4 | ෪<br>111E4  | ඨ<br>111F4  |
| 5 | ෫<br>111E5  |    |
| 6 | ෬<br>111E6  |    |
| 7 | ෭<br>111E7  |   |
| 8 | ෮<br>111E8  |  |
| 9 | ෯<br>111E9  |  |
| A | අ<br>111EA  |  |
| B | ඒ<br>111EB  |  |
| C | ඔ<br>111EC  |  |
| D | ඥ<br>111ED  |  |
| E | ඨ<br>111EE  |  |
| F | ඪ<br>111EF  |  |

**HISTORICAL DIGITS**

*These digits did not have a zero instead the numbers had signs for 10,20,30,40,50,60,70,80,90,100,1000. Sinhala Illakkam or Sinhala Archaic numbers were used for writing numbers prior to the fall of Kandyan Kingdom to British in 1815. These digits and numbers can be seen primarily in Royal documents and artifacts.*

- 111E1 ෧ SINHALA ARCHAIC DIGIT ONE
- 111E2 ෨ SINHALA ARCHAIC DIGIT TWO
- 111E3 ෩ SINHALA ARCHAIC DIGIT THREE
- 111E4 ෪ SINHALA ARCHAIC DIGIT FOUR
- 111E5 ෫ SINHALA ARCHAIC DIGIT FIVE
- 111E6 ෬ SINHALA ARCHAIC DIGIT SIX
- 111E7 ෭ SINHALA ARCHAIC DIGIT SEVEN
- 111E8 ෮ SINHALA ARCHAIC DIGIT EIGHT
- 111E9 ෯ SINHALA ARCHAIC DIGIT NINE

**HISTORICAL NUMBERS**

- 111EA ෧෦ SINHALA ARCHAIC NUMBER TEN
- 111EB ෨෦ SINHALA ARCHAIC NUMBER TWENTY
- 111EC ෩෦ SINHALA ARCHAIC NUMBER THIRTY
- 111ED ෪෦ SINHALA ARCHAIC NUMBER FORTY
- 111EE ෫෦ SINHALA ARCHAIC NUMBER FIFTY
- 111EF ෬෦ SINHALA ARCHAIC NUMBER SIXTY
- 111F0 ෭෦ SINHALA ARCHAIC NUMBER SEVENTY
- 111F1 ෮෦ SINHALA ARCHAIC NUMBER EIGHTY
- 111F2 ෯෦ SINHALA ARCHAIC NUMBER NINETY
- 111F3 ෧෦෦ SINHALA ARCHAIC NUMBER ONE HUNDRED
- 111F4 ෧෦෦෦ SINHALA ARCHAIC NUMBER ONE THOUSAND

## 1. References

Harsha Wijayawardhana, Numerations in the Sinhala Language, Information and Communication Technology Agency of Sri Lanka (ICTA), October 2009.

De Silva W.A., from 'Catalogue of Palm leaf manuscripts in the library of Colombo Museum', Volume I, Government Printer, 1938

Epa Panchanga (Epa Ephemeris), Epa Printers, 2007

Everson, Michael, ISO/IEC JTC1/SC2/WG2 N3195R, Proposal to add archaic numbers for Sinhala to the BMP of UCS, 2007

Gunasekera, Abraham Mendis, A Comprehensive Grammar of Sinhalese Language, Sri Lanka Sahitya Mandalaya (Academy of Letters, 144-150, 1891.

Hettigoda, Hendrick De Silva, Life and Planets, Vishwa Lekha, Sarvodaya, 34 -36, 1987 (First print in 1967)

Kularathne P.D.S, Article on Numeration, Sinhala Encyclopedia, 1967

**අභිලේඛන, සමරු පොත් පෙළ, දෙවන වෙළුම, ප්‍රධාන සංස්කාරක - පණ්ඩිත ආචාර්ය නන්දදේව විජේසේකර, පුරාවිද්‍යා දෙපාර්තමේන්තුව, 1990**

ගුණසේකර ඒබ්‍රහැම් මෙන්දිස් ග්‍රන්ථාවලිය, කතෘ - (පිටු අංක 3 සිට 10 දක්වා)



Figure 1: Lith Illakkam or Sinhala Astrological numbers found in an old Astrological Ola book (Kandy Museum)

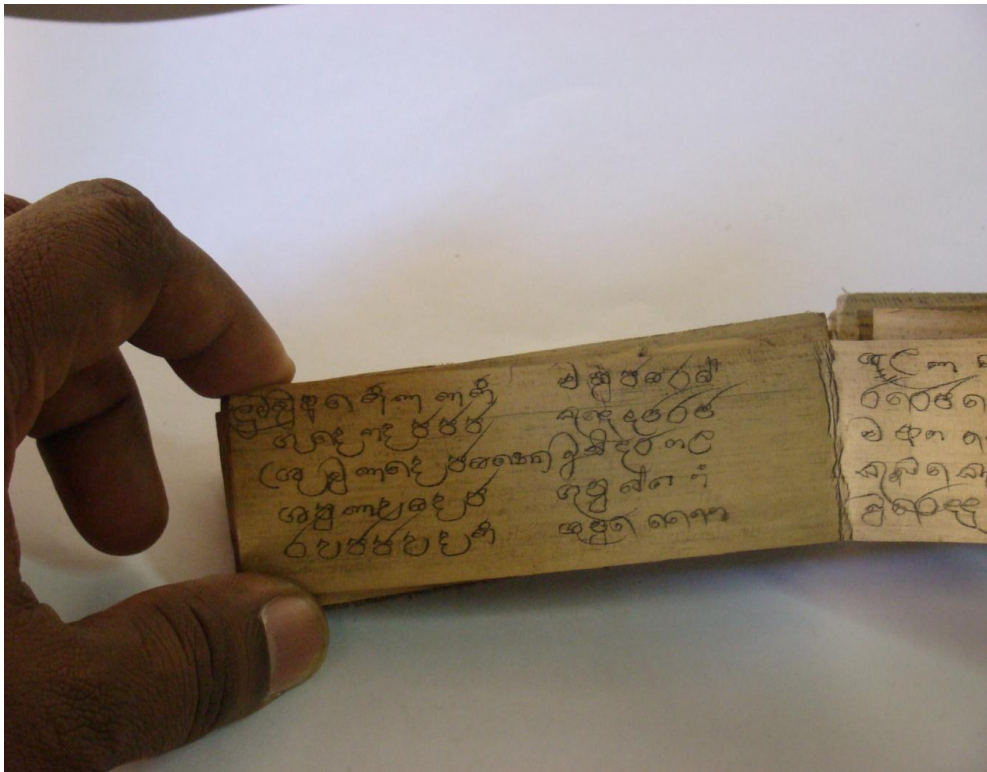


Figure :2 Planetary Positions depicted with Sinhala Lith Illakkam- Kandy Museum



Figure 3: 20th century horoscope with older version of Sinhala Lith Illakkam or Sinhala Astrological numbers



Figure 4: Ola book with page numbering in Sinhala Lith Illakkam or Sinhala Astrological numbers

**ISO/IEC JTC 1/SC 2/WG 2  
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS  
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646<sup>1</sup>**

**Please fill all the sections A, B and C below.**

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html>  
for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

**A. Administrative**

|   |   |  |            |
|---|---|--|------------|
| <b>1. Title:</b>  | <b>Proposal to include Sinhala Numerals to the BMP and SMP of the UCS</b> |  |            |
| <b>2. Requester's name:</b>   | <i>Sri Lanka Standards Institution (SLSI)</i>                             |  |            |
| <b>3. Requester type (Member body/Liaison/Individual contribution):</b> | <i>Member and National Body</i>   |  |            |
| <b>4. Submission date:</b>  | <i>2010-08-19</i>   |  |            |
| <b>5. Requester's reference (if applicable):</b>                        |   |  |            |
| <b>6. Choose one of the following:</b>                                  |   |  |            |
| This is a complete proposal:  |   |  | <i>Yes</i> |
| (or) More information will be provided later:                           |   |  | <i>No</i>  |

**B. Technical – General**

|   |  |                                    |            |
|---|--|------------------------------------|------------|
| <b>1. Choose one of the following:</b>  |  |                                    |            |
| a. This proposal is for a new script (set of characters):   |  |                                    | <i>No</i>  |
| Proposed name of script:  |  |                                    |            |
| b. The proposal is for addition of character(s) to an existing block:   |  |                                    | <i>Yes</i> |
| Name of the existing block:   | <i>Sinhala</i>   |                                    |            |
| <b>2. Number of characters in proposal:</b>   |  |                                    | <i>30</i>  |
| <b>3. Proposed category (select one from below - see section 2.2 of P&amp;P document):</b>  |  |                                    |            |
| A-Contemporary  | B.1-Specialized (small collection)   | B.2-Specialized (large collection) | <i>A</i>   |
| C-Major extinct   | D-Attested extinct   | E-Minor extinct                    |            |
| F-Archaic Hieroglyphic or Ideographic   | G-Obscure or questionable usage symbols  |                                    |            |
| <b>4. Is a repertoire including character names provided?</b>   |  |                                    |            |
| a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?   |  |                                    | <i>Yes</i> |
| b. Are the character shapes attached in a legible form suitable for review?   |  |                                    | <i>Yes</i> |
| <b>5. Fonts related:</b>  |  |                                    |            |
| a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?   | <i>Harsha Wijayawardhana, University of Colombo School of Computing (UCSC)</i>   |                                    |            |
| b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):  | <i>Harsha Wijayawardhana, University of Colombo School of Computing (UCSC), 35 Reid Ave Colombo 7<br/>Email: <a href="mailto:ghw@ucsc.cmb.ac.lk">ghw@ucsc.cmb.ac.lk</a>, <a href="mailto:wijayawardhana@gmail.com">wijayawardhana@gmail.com</a>, Web: <a href="http://www.ucsc.cmb.ac.lk/sdu">www.ucsc.cmb.ac.lk/sdu</a></i> |                                    |            |
| <b>6. References:</b>   |  |                                    |            |
| a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?   |  |                                    | <i>Yes</i> |
| b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?  |  |                                    | <i>Yes</i> |
| <b>7. Special encoding issues:</b>  |  |                                    |            |
| 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)? |  |                                    | <i>Yes</i> |

**8. Additional Information:**  
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 <http://www.unicode.org/Public/UNIDATA/UCD.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

<sup>1</sup> Form number: N3702-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11)

### C. Technical - Justification

|   |   |
|---|---|
| <p>1. Has this proposal for addition of character(s) been submitted before?<br/>If YES explain</p>  | <p><i>Yes for Sinhala Illakkam or Sinhala Archaic Numbers but not for Sinhala Lith Illakkam or Sinhala Astrological Numbers. See N3195R</i></p>   |
| <p>2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?<br/>If YES, with whom?</p>     | <p>Yes</p>  |
|   | <ul style="list-style-type: none"> <li>▪ <i>Local Language Group of the Information and Communication Technology Agency (ICTA), which consists of Linguists, Historians, Computer Experts etc.</i></li> <li>▪ <i>Research findings of Mr. Harsha Wijayawardhana were presented at National Archaeological Symposium held in 2009 in Colombo.</i></li> <li>▪ <i>Presentation made to Stakeholders comprising the National Museum, Royal Asiatic Society, Sri Lanka Standards Institution, Universities (including the Postgraduate Institute of Archaeology).</i></li> <li>▪ <i>Discussions held with Sigiriya Museum (National heritage site) and with astrologers, who use Sinhala numerals.</i></li> <li>▪ <i>Presentation made to the Sectoral Committee on IT of the Sri Lanka Standards Institution (SLSI). The Committee has approved in principle to add the numerals to the Sri Lanka Sinhala Character Code for Information Interchange, SLS 1134 : 2004.</i></li> </ul> |
| <p>If YES, available relevant documents:</p>  | <ul style="list-style-type: none"> <li>▪ <i>Mr. Harsha Wijayawardhana published his findings in a book titled "Numerations in the Sinhala Language".</i></li> <li>▪ <i>Proceedings of the National Archaeological Symposium, Volume II, 7<sup>th</sup> and 8<sup>th</sup> July 2009, Department of Archeology, Colombo 07.</i></li> <li>▪ <i>Minutes of the Sectoral Committee on IT, Sri Lanka Standards Institution (SLSI) May 2010.</i></li> </ul>   |
| <p>3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?</p> <p>Reference:</p> | <p><i>Sinhala Archaic numbers have historical use in the Sinhala script. Sinhala Astrological numbers are still use for writing Horoscopes by Astrologers</i></p> <p><i>Reference is provided in the reference list</i></p>   |
| <p>4. The context of use for the proposed characters (type of use; common or rare)</p> <p>Reference:</p>  | <p><i>Sinhala Archaic Numbers are rare but Sinhala Astrological numbers are still used</i></p> <p><i>Reference is provided in the reference list</i></p>  |
| <p>5. Are the proposed characters in current use by the user community?</p> <p>If YES, where? Reference:</p>  | <p>Yes.<br/>Astrologers in Sri Lanka still use Sinhala Astrological numbers which are known as Lith Illakkam in Sinhala</p> <p><i>Sinhala Archaic numbers are used for historical documents but</i></p>   |

*Astrological numbers are used for writing horoscopes even to the late 20<sup>th</sup> century. Before 1815, Astrological numbers were used for page numbering and calculations etc.*

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

*Astrological numbers are in the BMP and Sinhala Illakkam or Sinhala Archaic Numbers are in the SMP*

If YES, is a rationale provided?

If YES, reference:

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

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

*No*

If YES, is a rationale for its inclusion provided?

If YES, reference:

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

*No*

If YES, is a rationale for its inclusion provided?

If YES, reference:

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

*No*

If YES, is a rationale for its inclusion provided?

If YES, reference:

11. Does the proposal include use of combining characters and/or use of composite sequences?

*No*

If YES, is a rationale for such use provided?

If YES, reference:

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

If YES, reference:

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

*No*

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

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

*No*

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

If YES, reference:



