L2/05-087

Title:	**Correcting Malayalam Digit Zero glyph (U+0D66), and, adding Malayalam numerics (U+0D70, U+0D71, U+0D72)**
Author:	**Dr. Naga Ganesan, Houston, Texas**
Date:	**April 9, 2005**

1.0 Correct Glyph for Malayalam Digit Zero (U+0D66):

The glyph for Malayalam digit zero has not been given correctly in the Unicode code chart for Malayalam script. The glyph given at U+0D66 is in fact a representation of the fraction (0.5) which compares visually quite well with the parallel glyphs for fraction (0.5) in the neighboring Dravidian language scripts such as Tamil. Quite unlike the northern Devanagari script, Malayalam and Tamil scripts in general move away from conjunct formations. Additionally, the historic number systems of both Tamil and Malayalam are very similar.

Varghese Chacko, a native Malayalam speaker, has already presented a bug report to Unicode. Reference: http://www.unicode.org/~emuller/iwg/p29/
[Begin Quote]
The glyph given at location 0x0D66 is not '0'(Zero) as said in your documents (U0D00.pdf) The symbol you have given is the glyph for '1/2'. Malayalam has its own glyph or '3/4','1/2','1/4','1/8' etc... and the glyph for '0'(Zero) in Malayalam is similar to one given at 0x0D02.
[End Quote]

Tamil and Malayalam digit zero were introduced only in the nineteenth century. The correct glyph for Malayalam digit zero is a circle/oval (0) shape just as in any other Indic language script. A detailed Malayalam-to-Roman script transliteration document with diacritical marks has been developed by Dr. Thomas Pedersen, and the correct glyph of Malayalam digit zero in accordance with all Indic zero glyph is given by Pedersen:
http://transliteration.eki.ee/pdf/Malayalam.pdf

Table 1 in the following page provides the graphic shapes of Indic and other numerals. It is taken from the external source document on Indian numbers:
http://www.omniglot.com/language/numerals2.htm

Hence, it is recommended that Unicode code chart glyph for Malayalam digit zero (U+0D66) be changed to a circular form (0) which will be in agreement with the digit zero used in all Indic languages (eg., U+0BE6).

Table 1: Numbers in Malayalam (Indic) and world scripts
(Source: http://www.omniglot.com/language/numerals2.htm)

	0	1	2	3	4	5	6	7	8	9	10	100	1000	10000
Arabic	٠	١	٢	٣	٤	٥	٦	٧	٨	٩				
Bengali	০	১	২	৩	৪	৫	৬	৭	৮	৯				
Chinese (simple numerals)	〇	一	二	三	四	五	六	七	八	九	十	百	千	万
Chinese (complex numerals)	零	壹	貳	參	肆	伍	陸	柒	捌	玖	拾	佰	仟	萬
Chinese 花碼 (huā mă)	〇	〡	〢	〣	〤	〥	〦	〧	〨	〩				
Devanagari	०	१	२	३	४	५	६	७	८	९				
Ethiopic		፩	፪	፫	፬	፭	፮	፯	፰	፱	፲	፻		፼
Gujarati	૦	૧	૨	૩	૪	૫	૬	૭	૮	૯				
Gurmukhi	੦	੧	੨	੩	੪	੫	੬	੭	੮	੯				
Kannada	೦	೧	೨	೩	೪	೫	೬	೭	೮	೯				
Khmer	០	១	២	៣	៤	៥	៦	៧	៨	៩				
Lao	໐	໑	໒	໓	໔	໕	໖	໗	໘	໙				
Limbu	᥆	᥇	᥈	᥉	᥊	᥋	᥌	᥍	᥎	᥏				
Malayalam	൦	൧	൨	൩	൪	൫	൬	൭	൮	൯				
Mongolian	᠐	᠑	᠒	᠓	᠔	᠕	᠖	᠗	᠘	᠙				
Myanmar	၀	၁	၂	၃	၄	၅	၆	၇	၈	၉				
Oriya	୦	୧	୨	୩	୪	୫	୬	୭	୮	୯				
Tamil		௧	௨	௩	௪	௫	௬	௭	௮	௯	௰	௱	௲	
Telugu	౦	౧	౨	౩	౪	౫	౬	౭	౮	౯				
Thai	๐	๑	๒	๓	๔	๕	๖	๗	๘	๙				
Tibetan	༠	༡	༢	༣	༤	༥	༦	༧	༨	༩				
Urdu	۰	۱	۲	۳	۴	۵	۶	۷	۸	۹				

Note: (1) The digit zero including Malayalam is a glyph 0 as shown in this Table.
(2) The digit zero is added in March, 2005 to Tamil code chart in Unicode 4.1.

2.0 Malayalam Numerics 10 (U+0D70), 100 (U+0D71), 1000 (U+0D72):

In both Malayalam and Tamil scripts, digit zero is a modern innovation introduced in the nineteenth century. Like its closely allied Tamil script, Malayalam also employed numeric signs for 10, 100 and 1000 to write numbers larger than nine. Tamil and Roman zero-less numerical systems are compared by Michael Kaplan:
http://blogs.msdn.com/michkap/archive/2005/01/24/359347.aspx

Georges Ifrah, The universal history of Numbers from prehistory to the invention of the computer, John Wiley, 2000. Page 373 (the scanned page attached at the end),
" Malayalam figures

These figures are used by the Dravidian people of Kerala state, on the ancient coast of Malabar, in the southwest of India. They have the same name as the form of writing used in the area.

Like the Tamils, the people of Kerala did not use zero in their notation system for many centuries: Malayalam figures are not based on the place-value system, and there are specific figures for 10, 100 and 1,000. It was only since the middle of the nineteenth century, under the influence of Europe, that zero was introduced and combined with the symbols for the nine units according to the positional principle.

Thus the Tamil and Malayalam figures were the only ones in India that did not include zero and were not based on the positional principle relatively recently."

While Malayalam digit zero was introduced in 1850 CE or later, the Tamil Nadu government (then called Madras Presidency under the British colonial rule) introduced Tamil digit zero (U+0BE6) prior to 1820 CE. A citation from a 1825 CE school textbook is given at: http://www.geocities.com/thamizh@sbcglobal.net/tamil_zero.PDF

Like the Tamil script having numerics for 10, 100 and 1000 at U+0BF0, U+0BF1, U+0BF2 respectively, Malayalam historically also has numeric signs. Only with these numerics signs, large numbers greater than nine can be written in the traditional Malayalam script.

Hence it is recommended that Unicode allocates separate code points for the Malayalam numeric signs also.

Malayalam numerics
U+0D70 MALAYALAM NUMBER TEN
U+0D71 MALAYALAM NUMBER ONE HUNDRED
U+0D72 MALAYALAM NUMBER ONE THOUSAND

The code points are assigned in a manner parallel to the Tamil numerics code points. The shapes of the Malayalam numerics glyphs for U+0D70, U+0D71, U+0D72 must be as shown in page 335, Figure 23.21, Georges Ifrah, The universal history of Numbers from prehistory to the invention of the computer, John Wiley, 2000 which is enclosed in this proposal.

Author: Nagamanickam Ganesan, Ph.D.
16923 Sky Harbor Ct.
Friendswood, TX 77546, USA
281-648-8636 (Phone)
naga_ganesan@hotmail.com

Georges Ifrah,
The universal history of Numbers,
2000, pg. 334, John Wiley

DECISIVE FIRST STEP: THE PRINCIPLE OF POSITION

In order to achieve a system as ingenious as our own, it is first necessary to discover the *principle of position*. According to this, the value of a figure varies according to the position in which it occurs, in the representation of a number. In our modern decimal notation, a "3" has value 3 units, 3 tens or 3 hundreds depending on whether it is in the first, second or third position. To write seven thousand, six hundred and fifty-nine, all we have to do is to write down the figures 7, 6, 5, and 9 in that order, since according to the rule the representation 7,659 denotes the value

$$7 \times 1{,}000 + 6 \times 100 + 5 \times 10 + 9.$$

Because of this fundamental convention, only the coefficients of the powers of the base, into which the number has been decomposed, need appear.

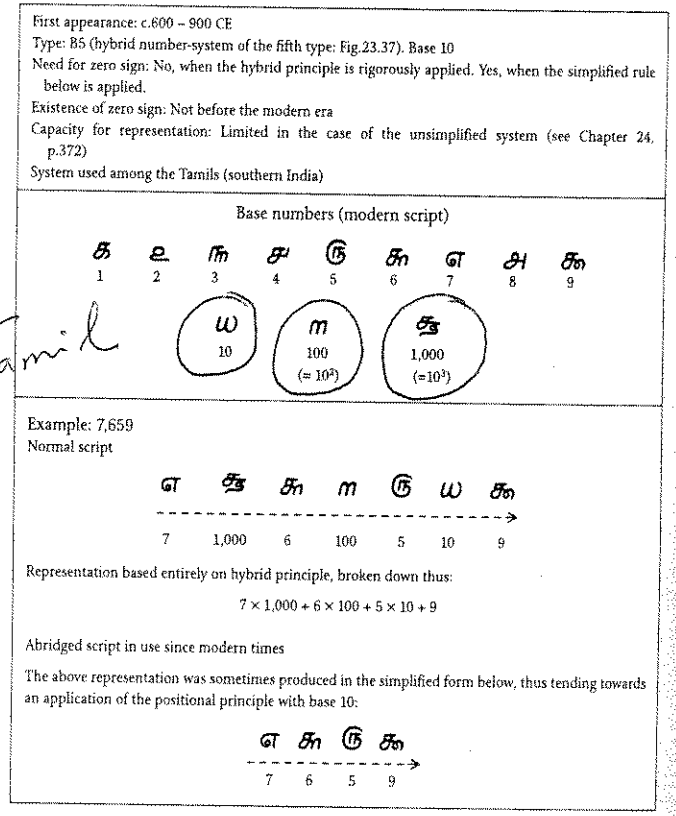
This, therefore, is the principle of position. Apparently as simple as Columbus's egg; but it had to be thought of in the first place!

Nowadays, this principle seems to us to have such an obvious simplicity that we forget how the human race has stammered, hesitated and groped through thousands of years before discovering it, and that civilisations as advanced as the Greek and the Egyptian completely failed to notice it.

SYSTEMS WHICH COULD HAVE BEEN POSITIONAL

For all that, even in the earliest times a goodly number of different number-systems could have led on to the discovery of the principle of position.

Consider for example the Tamil and Malayalam systems from south India. According to the hybrid principle, the figure representing the number of tens was placed to the left of the symbol for 10, the one representing the number of hundreds to the left of the symbols for 100, and so on (Fig. 23.20 and 23.21).

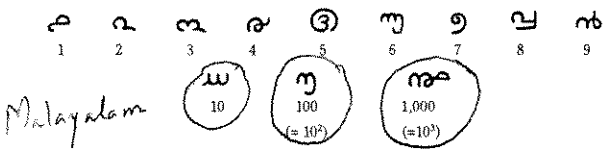


Tamil

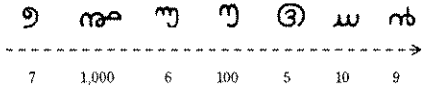
FIG. 23.20. *Tamil number-system*

First appearance: c.600 – 900 CE
Type: B5 (hybrid number-system of the fifth type: Fig. 23.37). Base 10
Need for zero sign: No, when the hybrid principle is rigorously applied. Yes, when the simplified rule below is applied.
Existence of zero sign: Not before the modern era
Capacity for representation: Limited in the case of the unsimplified system (see Chapter 24, p.373)
System used among the Malayalam (southern India, Malabar coast)

Base numbers (modern script)



Example: 7,659
Normal script



Representation based entirely on hybrid principle, broken down thus:

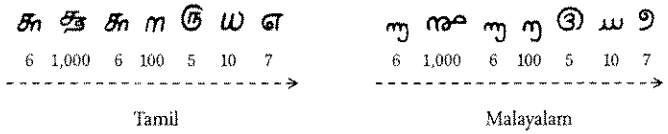
$$7 \times 1{,}000 + 6 \times 100 + 5 \times 10 + 9$$

Abridged script in use since modern times

The above representation was sometimes produced in the simplified form below, thus tending towards an application of the positional principle with base 10:

FIG. 23.21. *Malayalam number-system*

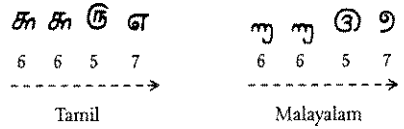
In this way, the number 6,657, for example, would usually be written as follows:



which corresponded to the decomposition

$$6 \times 1{,}000 + 6 \times 100 + 5 \times 10 + 7.$$

Now, when we look at certain Tamil or Malayalam writings, we find that the symbols for 10, 100, and 1,000 have in many cases been suppressed [L. Renou and J. Filliozat (1953)]. The number 6,657 would then appear in the abbreviated notation



The result of this simplification is that the figures 6, 6, 5, and 7 have been assigned values as follows:

- seven units to the figure 7 in the first place;
- five tens to the figure 5 in the second place;
- six hundreds to the figure 6 in the third place;
- six thousands to the figure 6 in the fourth place.

Thus the Tamil and Malayalam figures could be assigned values which depended on where they occurred in the representation of a number.

This remarkable potential for evolution towards a positional number-system is characteristic of hybrid numbering systems.

In such systems, in fact, the signs which indicate the powers of the base (10, 100, 1,000) are always written in the same order, either increasing or decreasing. Therefore it is natural that the people who used these systems would be led, for the sake of abbreviation, to suppress these signs leaving only the figures representing their coefficients.

This is what led certain Aramaic stone-cutters of the beginning of our era to sometimes leave out the sign for 100 in their numeric inscriptions.

G. Ifrah, 2000,

pg. 335

In other words, the first nine Tamil figures are from the same family as the other corresponding Indian numerical symbols, the difference lying in their style and their adaptation to the unique shape of Tamil writing.

Malayâlam figures

These figures are used by the Dravidian people of Kerala State, on the ancient coast of Malabar, in the southwest of India. They have the same name as the form of writing used in the area.

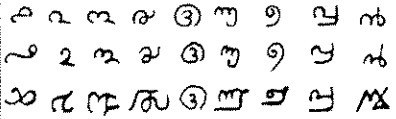
1	2	3	4	5	6	7	8	9	0	Ref.
										Drummond Frédéric, DCI Peet, J. Pihan Renouand Fillozat
Geographical area (Fig. 24. 27 and 24. 53): Used in the region stretching the length of the southeast coast of India, from Mangalore in the north to the southernmost point of India, and which is made up of a long coastal strip stretching from the coast of Malabar and by the Ghats encompassing the peaks of the Cardamoms.										

FIG. 24.19. *Current Malayâlam numerals*

Like the Tamils, the people of Kerala did not use zero in their notation system for many centuries: *Malayâlam* figures are not based on the place-value system, and there are specific figures for 10, 100 and 1,000. It was only since the middle of the nineteenth century, under the influence of Europe, that zero was introduced and combined with the symbols for the nine units according to the positional principle.

Thus the Tamil and *Malayâlam* figures were the only ones in India that did not include zero and were not based on the positional principle until relatively recently.

However, it should be noted that Tamil figures, a few centuries ago, before they evolved into their current forms, closely resembled their *Malayâlam* cousins which have conserved a style close to the original.

The graphical link with the numerical signs of other regions of India is more easily seen through examining the original appearance of the Tamil figures than through looking at their modern form (Fig. 24.17 and 24.19):

G. Ifrah, 2000

pg. 373