

Proposal to encode Tamil fractions and symbols

Shriramana Sharma, jamadagni-at-gmail-dot-com, India

2012-Jul-17

§1. Thanks

I owe much to everyone who helped with this proposal.

G Balachandran of Sri Lanka heavily contributed to this proposal by sharing the attestations and data which he had collected over five years back in researching this very same topic. Dr Jean-Luc Chevillard of France and Dr Elmar Kniprath of Germany provided the second majority of attestations. Dr Kalyanasundaram (Switzerland), Dr Jayabarathi (Malaysia), K Ramanraj (Chennai), Vijayaraghavan Vanbakkam (Germany), Dr Rajam (US), Dr Vijayavenugopal (Pondicherry), Mani Manivannan (Chennai/US) and A E Elangovan (Chennai) also helped with attestations. Various members of the CTamil list also contributed to related discussions.

Vinodh Rajan of Chennai is my personal sounding board for all my proposals and he contributes in too many ways for me to specify. Deborah Anderson continuously helped out with encouragement and enthusiasm-bolstering :-).

I express my sincere thanks to all these people and to anyone else not mentioned (my apologies!). Preparing this has been a lengthy journey, and you all helped me through!

§2. Introduction

Compared to other Indic scripts/regions, the Tamil-speaking region has employed a larger set of symbols for fractions and abbreviations. These symbols, especially the fractions, have been referred to in various documents already submitted to the UTC, including my Grantha proposal L2/09-372 (p 6). A comprehensive proposal has been desirable for quite some time for the addition of characters to Unicode to enable the textual representation of these rare heritage written forms which are to be found in old Tamil manuscripts and books.

Previously submitted documents proposing a subset of these characters are L2/09-376, L2/10-334R, L2/10-428 (all by the present author) and L2/10-408 (by Naga Ganesan for INFITT). Other feedback documents relevant to this issue are L2/09-398 (by Naga Ganesan for INFITT) and L2/09-416 (by ICTA Sri Lanka). Action item 125-A42 is also to be noted.

(For reference, the currently encoded Tamil symbols were proposed in L2/01-375R.)

In summary, the consensus is that all Tamil fractions should be encoded together, and any other attested symbols (apart from the ones already encoded) should also be encoded together with those fractions.

It has always been clear that the BMP Tamil block (despite many empty spaces) cannot contain the whole set of these additional characters. I had hence submitted L2/09-316 asking for a supplementary Tamil block to be allocated in the SMP roadmap for the encoding of these characters. Fulfillment of said request has been delayed pending a clear idea of the number of columns that would be required. That delay may now be cleared and a block be allocated by the name Tamil Supplement (after Latin-1 Supplement, etc).

This document proposes to encode sixty two (62) characters. Seven (7) characters related to existing BMP characters are proposed to be added to the Tamil BMP block. Even then, a block of four columns is required in the SMP. Already some potentials for future encoding (pending further research) are identified here. Hence such a block is justified.

It is suggested that the unallocated empty range 11FC0-11FFF at the far end of the current Brahmic scripts zone in the SMP be allocated for this block. It is likely that the remaining space in this zone would be filled up by archaic Brahmic scripts, and so isolating a supplementary block meant for fractions, symbols etc to the far end would make sense.

In the course of the research for this proposal, it has come to light that the glyphs for two existing encoded Tamil symbols need to be corrected to match the attestations. This has already been treated in my document L2/12-106 §2. It has also been discovered that the character names of two encoded written forms do not correspond with their usage as documented in authoritative academic sources. Hence some additional character encoding and cross-references are needed. All this is summarized in §5.

Further to be noted is that a set of characters has already been recognized by the Tamil Nadu Government (henceforth, GOTN) and documented in L2/10-318, esp on pp 15, 27, 28, 38, 39, 42 and 43. A review of this standard in the light of this proposal is provided in §6. It would be good if the GOTN could evaluate the contents of this (brief) review and incorporate the suggestions provided into their standard/GO.

It is of course understood that interested institutions like the GOTN, ICTA Sri Lanka and INFITT might like to provide feedback on this document. All constructive feedback is of course welcome in the interests of providing a proper representation of all attested Tamil-related written forms in Unicode.

With these preliminaries, I proceed to introduce the actual set of characters.

§3. List of characters

§3.1. Fractions

	Glyph	x/320	Value	Words	Tamil word
1.	முதி	1/320	1/320	one three-hundred-and-twentieth	muntiri
2.	ரி	2/320	1/160	one one-hundred-and-sixtieth	araikkāṇi
3.	ு	4/320	1/80	one eightieth	kāṇi
4.	ு	5/320	1/64	one sixty-fourth	kālvīcam
5.	சு	8/320	1/40	one fortieth	araimā
6.	சு	10/320	1/32	one thirty-second	aravīcam
7.	சு	12/320	3/80	three eightieths	mukkāṇi
8.	சு	15/320	3/64	three sixty-fourths	mukkālvīcam
9.	ப	16/320	1/20	one twentieth	mā
10.	ப	20/320	1/16	one sixteenth	vīcam / mākāṇi
11.	பி	32/320	1/10	one tenth	irumā
12.	பு	40/320	1/8	one eighth	araikkāl
13.	பு	48/320	3/20	three twentieths	mummā
14.	பு	60/320	3/16	three sixteenths	mūvīcam / mummāmukkāṇi
15.	பு	64/320	1/5	one fifth	nāṅkumā
16.	பு	80/320	1/4	one quarter	kāl
17.	பு	160/320	1/2	one half	arai
18.	பு	240/320	3/4	three quarters	mukkāl

There is also the fractions qualifier kīl:

19.	கீ	1/320	× 1/320	multiply by 1/320	kīl
-----	----	-------	---------	-------------------	-----

... which when prefixed to a fraction, reduces (downscales) the value of the fraction by a scale of 1/320 (the value of the lowest fraction muntiri).

§3.2. Measures of grain

	Glyph	Meaning	Tamil word
1.	ஐ	1 grain of paddy	nel
2.	ஐ௬	360 nel	cuvaṭu
3.	ஐ	5 cuvaṭu	āḷakku
4.	ஐ	2 āḷakku	uḷakku
5.	ஐரி	2 uḷakku	uri
6.	ஐ	3 uḷakku	mūvuḷakku
7.	ஐ (ALREADY ENCODED)	2 uri (4 uḷakku)	nāḷi / paṭi
8.	ஐ	8 nāḷi / paṭi	kurun̄i / marakkāl
9.	ஐ	2 kurun̄i	patakkū
10.	ஐ	3 kurun̄i	mukkurun̄i
11.	ஐ	2 patakkū (4 kurun̄i)	tūṇi
12.	ஐ	3 tūṇi	kalam

Simple ligatures of numbers with these measures are seen. They are not separately encoded as they can be easily provided by fonts by glyph substitution:

patakkū = 2 kurun̄i = ஐஐ → ஐஐ

2 tūṇi = ஐஐ → ஐஐ*

1 kalam = கஐ → கஐ**

* This is also confusable with the fraction muntiri ஐஐ.

** It should be noted that when there is only one full kalam (i.e. there is only one kalam at all or one kalam and other lesser measures, for a total of less than two kalam-s), it is always written as the ligature கஐ and never as just the symbol for kalam i.e. ஐ. Thus while other ligatures are optional, a good font *must* replace க + ஐ → கஐ for “one kalam”.

§3.3. Currency

	Glyph	Meaning	Tamil word
1.	பசு	small currency unit	paicā
2.	அணு	4 paisā	aṇā
3.	ரூபி (ALREADY ENCODED)	16 aṇā	rūpāy
4.	கூ	small currency unit	kācu
5.	பு	80 kācu	paṇam
6.	பை	10 paṇam	pon
7.	வாண	42 paṇam	varākaṇ
8.	ரிங்கி	ringgit	riṅkiṭ

§3.4. Miscellaneous

§3.4.1. Weight, length and area

	Glyph	Meaning	Tamil word
1.	பாரம்	≅ 500 pounds of weight	pāram
2.	கேஜம்	≅ yard	kejam
3.	குலி	16 sq. kejam	kulī
4.	வெலி	2000 kulī	vēli

§3.4.2. Agriculture

5.	நாணய	wet cultivation	nancey
6.	பண்ணை	dry cultivation	punṇey
7.	நிலம்	land	nilam
8.	அளம்	salt pan	aḷam

§3.4.3. Clerical

9.	வரவு	credit	varavu
10.	நா	number	eṇ
11.	நிலை	current	nāḷatu

12.	ஈ	... and odd	cilvāṇam / cillarai
13.	பீ	... having been spent	pōka
14.	அ	total	āka

§3.4.4. Other abbreviations

15.	வ	in the charge of	vacam
16.	மு	first	mutal
17.	மு	et cetera (in a series)	mutaliya
18.	வ	et cetera (of a kind)	vakaiyarā
19.	ஈ	long-lived	cirañcīvi
20.	பீ	a title	piḷḷai
21.	ரீ	respected	rāja

§3.4.5. Abbreviated word-endings

22.	டு	... for *	...kku
23.	யு		...yum
24.	வு		...vum

* usually after numerals in the sense of “at the rate of”

§3.4.6. Punctuation

25.	௯	end of text
-----	---	-------------

§3.5. Summary

Listed above are 19 symbols for fractions, 12 for measures, 8 for currency and 25 miscellaneous for a total of 64. Of these, two characters are already encoded and are only included in the above list for sake of completeness of the list in relation to other characters. Thus, 62 characters need to be encoded.

The shape and usage of these characters are immediately discussed in §4. Their relation to existing characters is considered in §5. The TN Govt standard which includes some of these characters is considered in §6. Attestations are collected in §7.

§4. Details of shape and usage

§4.1. Loss of Tamil numerals and symbols from vogue

At the outset of this proposal, we mentioned the fact that it is the Tamil script which developed so many different fractions and symbols compared to all other Indic scripts. Sadly, most of these characters are not in use today and most Tamilians today are totally unaware of these old written forms.

It is obvious that the minor fractions and measures went into disuse due to the advent of the decimal system and old currency symbols were discontinued due to changes in the currency in vogue. However, even the rest of the Tamil numerals and the remaining abbreviations have gone out of use due to various social changes.

Thus these written forms are of interest mostly to those who work with old writings that use them, perhaps (hopefully) to digitize them as e-text. A small number of these characters are however still used in traditional contexts, especially such as marriage invitations written in Tamil and printed in a traditional format.

§4.2. Resolution of glyphic variants and variant representations

As a result of the absence of these characters from common use, many of them do not have a totally consistent written form that is uniform throughout all the (few) texts that use them. One finds attestation for similar *and* dissimilar alternative written forms for many of these characters. It is hence needed to sift through these alternates to arrive at a standard.

A careful examination of the evidence indicates that in a few cases, alternative representations developed naturally. On the other hand, in the case of many numerals including fractions a distinct tendency to give them letter-like forms is seen.

However in many other cases, specific written forms were forgotten over time and substituted by other similar shapes. On occasion, a distinct character has been substituted into the role of a similar/related character, and not always due to glyphic similarity. It has been necessary to weed away such substitutes to arrive at the true representation. In this, we have mostly benefited from a fortuitous gain of one original attestation of which later attested forms are clearly approximations due to the loss of printer's glyph repertoire.

In the few cases where there have been legitimate variants of some characters, we have chosen that written form as representative which we judged least confusable with the written forms of other similar characters. This applies to both cases of normal glyphic variants and to cases of divergent evolution from a common root as mentioned above.

§4.3. Introduction to category-wise discussion

In the following sub-sections we discuss the highlights of the various categories of proposed characters. We have also presented the reasons for our making specific choices as to standardized glyphs. During this discussion, repeated references will be made to the various attestations reproduced in §7. Other pertinent issues (such as disunification of characters, characters not proposed now and slight glyph changes required for existing characters) are discussed in later sub-sections of the current §4.

§4.4. Fractions (see §3.1)

The Tamil fractions are divided into two series: a series used for generic measurements and calculations and one used especially with money. In the general series, below the major fractions $3/4$, $1/2$ and $1/4$, a fraction $1/5$ nālumā is introduced, and is further subdivided down by the major fractions, i.e. $1/5 \times 3/4 = 3/20$ and so on, to go up to $1/320$. The fractions for money however do not introduce $1/5$, and it is $1/4$ that is subdivided as before to produce fractions with denominators as powers of two: $1/4 \times 3/4 = 3/16$ and so on. The lowest fraction in this series is $1/64$, which is the ratio of an old paisa to a rupee. (See also §4.7.)

In the general series probably the only ligature is the glyph for mukkāṇi:

$$\text{mukkāṇi } 3/80 = \text{araimā } 1/40 + \text{kāṇi } 1/80 = \text{சு௮} \rightarrow \text{சு௮}$$

Perhaps irumā ௨ which equals 2 mā-s (iru = 2) is also a ligature of 2 + mā ௨ + ௮.

However among “money-fractions” many glyphs are clearly ligatures with the generic Tamil abbreviation mark ௨ of glyphs of lesser fractions from the general series which bear to the former the ratio 4:5. Without these ligatures, these “money-fractions” would have to be written out as a sequence involving those lesser fractions, and the ligatures were probably a later development for convenience. This is illustrated as follows:

$$\begin{aligned} \text{kālvīcam } 1/64 &= \text{kāṇi } 1/80 + \text{muntiri } 1/320 = \text{௮௪௪} \rightarrow \text{௮} + \text{௨} = \text{௮௪} \\ \text{aravīcam } 1/32 &= \text{araimā } 1/40 + \text{araikkāṇi } 1/160 = \text{சு௮௪} \rightarrow \text{சு௮} + \text{௨} = \text{சு௮௪} \\ \text{mukkālvīcam } 3/64 &= \text{mukkāṇi } 3/80 + \text{araikkāṇi } 1/160 + \text{muntiri } 1/320 = \text{சு௮௪௪௪} \\ &\rightarrow \text{சு௮} + \text{௨} = \text{சு௮௪} \\ \text{mūvīcam } 3/16 &= \text{mummā } 3/20 + \text{mukkāṇi } 3/80 = \text{௪௮௪} \rightarrow \text{௪௮} + \text{௨} = \text{௪௮௪} \end{aligned}$$

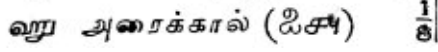

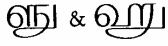

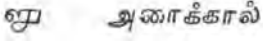

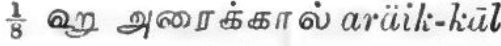
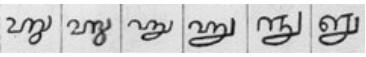
The notable exceptions to money-fractions being ligatures are vīcam $1/16$ ௮௪ and araikkāl $1/8$ ௪௪, which were presumably more often found in common usage and are even attested historically by Subramanian and Caa Ganesan.

Some sources like Beschi do not list the ligated money-fractions.

The present proposal integrates the two series of fractions into a single list in uniform ascending order to facilitate collation and comparisons.

§4.4.1. Significant glyphic variants

Of all fractions, araikāl has the most number of significant glyphic variants:

Cintamani En Suvadi / Kanakkadigaram 1958	:	
Kanakkadigaram 1958 (again)	:	
Gruenendahl	:	
Beythan	:	
Beschi	:	
Shuddhananda Bharati	:	
Wickremasinghe	:	
Caa Ganesan	:	

One notes that Caa Ganesan has effectively shown the evolution of the form shown by Beschi from the other forms. However, this form is confusable with the sign for the measure ālakku which is exclusively written as ளு. Therefore we have chosen the older form but included the angular feature (as attested by most sources), so: ளு for araikāl.

§4.4.2. Alternate representations

Ⓜ is attested as a glyph for arai $\frac{1}{2}$ by Subramanian, Burnell and Gruenendahl. Kanakkadigaram 1958 and Tamil Lexicon attest இ. Caa Ganesan shows how the latter derives from the former. However most sources (including those that show இ except Caa Ganesan) attest ள for arai. ள is easily construed as a simplified alternate evolution from the old Ⓜ. In the interest of following the most prevalent attestation and distinguishing from the Vowel Letter I இ, the glyph ள is chosen as representative.

The variant representations of the money fractions have been mentioned. Of these, the vīcam or mākāṇi $\frac{1}{16}$ bears special mention. It is often written as the sequence of mā $\frac{1}{20}$ + kāṇi $\frac{1}{80}$ i.e. பௗ. A ligated form of this viz பூ is also seen in Cintamani En Suvadi. Caa

Ganesan however shows that the historical version is ஸ, which is also attested by Kanakkadigaram 1958. While there is a resemblance of ஸ to டு, this is perhaps only fortuitous as the former is historical. Of these forms, ஸ is chosen as representative as it is the older and single-glyph representation.

§4.4.3. Errant attestations

Kanakkadigaram 1958 does not show the proper glyph for araivīcam $1/_{32}$ i.e. ஸ as seen in Cintamani En Suvadi but conflates it with the one for araimā $1/_{40}$ ஸ.

Burnell and Gruenendahl (who probably only derives from Burnell in this case) show டு for vīcam $1/_{16}$. However, this is probably an error because this is not corroborated by other sources. The glyph shown is in fact identical to that shown by Subramanian and Caa Ganesan for historical araikkāl $1/_{8}$. Burnell was an early author and presumably did not have as much material in evidence as later authors. He hence was probably mistaken as to the identity of டு.

§4.5. Lower fractions

The fraction multiplier kīl கீ is prefixed to fractions to downscale their value by a factor of $1/_{320}$. கீ is an abbreviation for the word கீழ் (low) which is sometimes written out in full or as the first letter alone (as in Shuddhananda Bharati). In effect, the kīl scale of numbers equates its onru i.e. 1 with muntiri i.e. $1/_{320}$ and the fractions apply to that.

Thus we have:

கீ க	=	kīl onru	=	$1/_{320} \times 1$	=	$1/_{320}$
கீ ஞ	=	kīl mukkāl	=	$1/_{320} \times 3/4$	=	$3/_{1280}$
கீ ற	=	kīl arai	=	$1/_{320} \times 1/2$	=	$1/_{640}$
கீ வ	=	kīl kāl	=	$1/_{320} \times 1/4$	=	$1/_{1280}$
...						
கீ ஞ	=	kīl kāṇi	=	$1/_{320} \times 1/80$	=	$1/_{25600}$
கீ ற	=	kīl araikkāṇi	=	$1/_{320} \times 1/160$	=	$1/_{51200}$
கீ டு	=	kīl muntiri	=	$1/_{320} \times 1/320$	=	$1/_{102400}$

There are fraction scales even below this but curiously, the continuation is not uniform. These are discussed below for the record.

◊ The value of $kīl$ muntiri is defined to be $10\frac{1}{2}$ on the “immi” scale. The value of “immi” one is therefore:

$$\text{immi onru} = kīl \text{ muntiri} / 10\frac{1}{2} = \frac{1}{102400 \times 10.5} = \frac{1}{1075200}$$

One then has immi mukkāḷ, immi arai and so on upto immi muntiri, which is equal to:

$$\text{immi muntiri} = \text{immi onru} / 320 = \frac{1}{1075200 \times 320} = \frac{1}{344064000}$$

◊ The value of immi muntiri is defined to be 3 on the “nuṇmai” scale. The value of “nuṇmai” one is therefore:

$$\text{nuṇmai onru} = \text{immi muntiri} / 3 = \frac{1}{344064000 \times 3} = \frac{1}{1032192000}$$

One then has nuṇmai mukkāḷ, nuṇmai arai and so on upto nuṇmai muntiri, which equals:

$$\text{nuṇmai muntiri} = \text{nuṇmai onru} / 320 = \frac{1}{1032192000 \times 320} = \frac{1}{330301440000}$$

◊ The value of nuṇmai muntiri is defined to be $10\frac{3}{4}$ on the “ciṇṇa” scale. The value of “ciṇṇa” one is therefore:

$$\text{ciṇṇa onru} = \text{nuṇmai muntiri} / 10\frac{3}{4} = \frac{1}{330301440000 \times 10.75} = \frac{1}{3550740480000}$$

One then has ciṇṇa mukkāḷ, ciṇṇa arai and so on upto ciṇṇa muntiri, which equals:

$$\text{ciṇṇa muntiri} = \text{ciṇṇa onru} / 320 = \frac{1}{3550740480000 \times 320} = \frac{1}{1136236953600000}$$

◊ The above scales are provided as per Cintamani En Suvadi*.

The scales of fractions end here. Cintamani En Suvadi also elucidates that since “ciṇṇa” is the smallest scale of “piṇṇam”-s i.e. fractions, the term “ciṇṇa-piṇṇam” figuratively refers to very small parts, from which even the common man (with total ignorance of the etymology) corrupts this as “ciṇṇāpiṇṇam”.

It is unclear however as to what practical application these infinitesimal fractions could have had. Fractions on the $kīl$ scale are found in texts of mathematical instruction like the Kanakkadigaram, but neither attestations for actual usage of the “immi” “nuṇmai” or “ciṇṇa” fractions nor for multiplier symbols for these scales could be found.

* This text provides the values of the lower scales w.r.t. to the $kīl$ scale i.e. with $kīl$ 1 = normal $\frac{1}{320}$. I have provided the values in the normal scale and double-checked them. In doing so I find that this text has gone inexplicably wrong in providing the value of ciṇṇa muntiri even in the $kīl$ scale. The publisher in his preface comments on the errors (especially w.r.t. the symbols used) in previous publications and asserts that he has made special efforts towards correctness. However, I find in this case and that of the symbols for the measures patakku and irutūṇi (see §4.6.3) that he has sadly not exceeded the saying “to err is human”.

Shuddhananda Bharati mentions a fraction $\frac{1}{1838400}$ aticāram which neither has an attested symbol nor fits within any of the above scales.

§4.6. Measures of grain (see §3.2)

The Tamil measures of grain are pretty straightforward, although certainly not decimal! Kanakkadigaram 1958, Cintamani En Suvadi and Beschi provide the list of various measures and their progression from “nel” (one grain) to “kalam” (and sometimes beyond).

Of these measures, the one evident ligature is uri ஊரி , which derives from the written-out word உரி run together. ஈத for mūvuḷakku (equal to 3 uḷakku -s) is also a possible ligature of mūru i.e. the digit ந and uḷakku ஊ .

The sign for the measure kalam is கா . There is a custom with regard to its use. Normally, the digit one ஈ is not prefixed to the other measures. For example, when there is only one marakkāl/kuruṇi , it is written as just நா . However, when there is only one kalam, i.e. the total volume is only one kalam or one kalam and odd for a total of less than two kalam-s, then it is always presented as the ligature of one + kalam = $\text{ஈ} + \text{கா} = \text{காஈ}$ and never as just கா . For example, when there are 13 marakkāl -s, where 12 marakkāl -s make 1 kalam, the written form is 1 kalam + $\text{marakkāl} = \text{காநா}$ and never காநா . Beyond this, i.e. for two kalam-s or more, it is obviously just கா prefixed by the numerals, such as உகா , நகா , and so on. This distinction must be kept in mind when one variously sees the glyph for kalam as காஈ and கா in the attestations. The above usage distinction is only explained by Beschi.

There are also other ligatures seen in the printings, such as 2 marakkāl/kuruṇi written as $\text{உ} + \text{நா} = \text{உநா}$ and 2 tūṇi written as $\text{உ} + \text{த} = \text{உத}$. However, they are optional. We have already briefly mentioned these cases along with the above kalam issue in §3.2.

Note that while some sources list irutūṇi (i.e. 2 tūṇi -s) as a distinct measure, there is no unique glyph provided for it. Most of these sources merely show the ligature of 2 + tūṇi i.e. உத . Beschi writes this merely as the sequence உத . We shall also show that the glyph seen for this in Cintamani En Suvadi is faulty. Thus no distinct character is encoded for this.

§4.6.1. Significant glyphic variants

The only significant case of glyphic variants is that of mukkuruṇi நு which is also seen as நு . We have chosen the first glyph to distinguish the measure from the fraction mummā

³/₂₀. While that fraction is also attested by Gruenendahl to be written as ᱠᱚ , this is not seen in any of the other sources, and Beschi contrasts mukkurūṇi ᱠᱚ with mummā ᱠᱚ and we have chosen to follow him in this case. (Gruenendahl does not list mukkurūṇi at all.)

A case, not of proper glyphic variants but of difficulty in identifying the proper glyph, is in the case of nel, one grain, the least measure of volume. The finalized glyph is ᱠᱚ as attested by Wickremasinghe (which source I obtained only relatively late in my research). Others have variously approximated it by a trimmed-off JA (Pope), a full JA ᱠᱚ (Winslow and Gruenendahl), different handwritten glyphs (Gruenendahl and Kanita Nuul) or not used a symbol at all (Cintamani En Suvadi, Kanakkadigaram 1958).

§4.6.2. Alternate representations

The ligature of 2 marakkāl/kuruṇi written as $\text{ᱠᱚ} + \text{ᱠᱚ} = \text{ᱠᱚᱠᱚ}$ is the only shape shown by Beschi for patakku, which is attested as ᱠᱚᱠᱚ by Kanakkadigaram 1958 and as a similar but somewhat simplified glyph (albeit confusable with araikkāl ᱠᱚᱠᱚ) by Winslow and Gruenendahl. The shape shown by Kanakkadigaram is chosen as the representative form, as the ligature ᱠᱚᱠᱚ may be obtained by glyph substitution in a supporting font.

Further for this same measure patakku, Cintamani En Suvadi shows the alternate shape ᱠᱚ and calls to our attention in a footnote that the (similar?) shape ᱠᱚ is used for the fraction irumā ¹/₁₀. One also notes that the shape ᱠᱚ shown by this text for mukkurūṇi is also the sign for the fraction mummā ³/₂₀, and patakku : mukkurūṇi :: irumā : mummā :: 2 : 3. Perhaps a shape similar to the fraction for irumā was used occasionally for patakku. However this shape is not seen in any of the other sources. We will also presently note that the Cintamani En Suvadi probably was at error for even the other shape for patakku. Therefore, we stay with the Kanakkadigaram 1958 glyph for patakku.

§4.6.3. Errant attestations

Cintamani En Suvadi has apparently incorrectly interchanged the glyphs for patakku and irutūṇi and showed them as ᱠᱚ and ᱠᱚ respectively. Most sources show them the other way around, and in fact given that tūṇi is ᱠᱚ , it is only natural that irutūṇi is $\text{ᱠᱚ} + \text{ᱠᱚ} = \text{ᱠᱚᱠᱚ}$. Thus we conclude that there is no distinct shape for irutūṇi and do not distinctly encode it.

§4.7. Currency (see §3.3)

The rupee / ana / paisa system is the commonly known currency system of India. (At least until my childhood I heard the word ana used though not in its exact original value.) An alternate system (which the Kanita Nuul mentions as being in vogue “in some regions”) is that of panam, pon and varagan. Unsurprisingly, neither of these were decimal, although the rupee / paisa system (abolishing the ana) was *made* decimal in the last century.

The relationships between these denominations and many other minor (intermediary) denominations is listed out in detail in the various sources. Only a few of these denominations have actual attested symbols and are proposed for encoding. Perhaps further research might turn out some more.

The one particular issue that deserves mention is that of pai / paisa. The symbol attested is ூ. which is clearly a derivation from Tamil pa ூ and ca ூ. Gruenendahl labels this “paisa or pie”. Winslow labels this “paisa” in Tamil but “pie” in English. Arden does not show ூ but only the Tamil syllable ூ (which literally reads “pai”) for “pie”. He does not mention any “paisa”. Pope is like Winslow in the symbol and nomenclature. Curiously, the AES edition of his text gives the written form as ூ (it is unclear what the ூ is in aid of). Wickremasinghe is the same with the names, but gives the written form as only ூ.

We note that many sources seem to imply that paisa/paicā and pai are one and the same denomination. Others* clarify that 1 rupee = 16 annas = 64 paise = 192 pies. Cintamani En Suvadi does not name paisa but names the $1/192$ part of a rupee as *tampiṭi* and *calli*. Paisa would be the *kāl aṇā* (i.e. $1/4$ ana) mentioned here. (Note that *tampiṭi* and *callikkācu* are both still words in common Tamil usage signifying something of very small value.)

Whatever be the values, it is clear that the syllable ூ (and the probably erroneous ூ) is not a candidate for encoding. Neither is the letter ூ seen in some sources for aṇā instead of ூ, or the syllable ூ seen for rupee instead of (the already encoded) ூ. The symbol ூ on the other hand should be encoded, and given its glyphic derivation from Tamil pa ூ and ca ூ it is named for paisa/paicā and not pie, whether the two be one or not.

* like http://en.wikipedia.org/wiki/Decimalisation#Rupee-anna-paisa-pie_conversion
<http://indiacoinecollections.blogspot.in/2011/09/pie-pice-naya-paisa-hindi-paisa-naya.html>
<http://stampsopakistan.tripod.com/glossary.htm> all retrieved 2012-Jul-09

In passing, one notes that some sources given a lesser known alternate glyph for rupee where the abbreviation mark ௨ is joined with the ரூ (rū, for rūpāy) below the base and not above (as currently shown in the code chart).

The ringgit sign ரீ is quite rare but attested to by a Malaysian Tamil scholar (Dr Jayabarathi) and recognized by others (via Muthu Nedumaran, email dated 2012-Feb-21).

§4.8. Miscellaneous symbols and abbreviations

§4.8.1. Measures of weight, length and area (see §3.4.1)

Cintamani En Suvadi, Kanakkadigaram 1958 and Kanita Nuul list in detail the various measures of weight, length and area used by the Tamils in the olden days. Other sources list only some of these. Very few of these have distinct attested symbols, though.

One measure of length the kejam (also kajam or kecam, but always pronounced gejam) which is equated to a yard (3 feet) is still in common usage today, especially in regard to sarees. Its sign கீ is however quite forgotten. (The half-measure of the kejam, the muḷam, is used for dhoti-s and is known even to the common flower-sellers who approximate it by the length of their upper arm. There is no symbol seen for this.)

One measure of area is the kuḷi (கூ), and it is equal to one square kōl (which is a measure of length with no attested symbol). However, there are (at least) two kōl-s. Cintamani En Suvadi equates a kōl with 10 feet which would make a kuḷi equal to 100 sq feet. Kanakkadigaram 1958 however says that 1 kuḷi is a square kōl with a 12 foot = 4 yard = 4 kejam kōl, making it equal to 144 sq feet = 16 sq yards = 16 sq kejam. Kanita Nuul calls this 4 kejam kōl a cemporkōl and affirms that 1 kuḷi is 144 sq feet = 16 sq kejam. (Another kōl called the ciṛukōl is equated to 2 muḷam = 1 kejam by Kanita Nuul but 12 muḷam by Kanakkadigaram 1958. It is not mentioned in Cintamani En Suvadi. It however neither has any attested symbol nor is it related in any specific way to the kuḷi.)

One notes that the kuḷi is translated as one square yard by Wickremasinghe, Pope, Arden, Winslow and Gruenendahl i.e. all our English sources. However, in deference to the natives' definition and considering their greater detail as seen in the native texts in the previous paragraph, we have annotated this character as 144 (and not 1) sq kejam.

In passing one notes that Arden shows a glyph for kuḷi in which the abbreviation mark ௨ is joined with the கு (ku, for kuḷi) below the base and not above as seen in most sources. (This is like for the rupee sign as noted before. Arden has most glyphs this way.)

Now considering higher measures of area than kuḷi, one finds that 100 kuḷi = 1 mā. 4 mā = 1 kāṇi (only mentioned in Cintamani En Suvadi). 2000 kuḷi = 20 mā = 5 kāṇi = 1 vēli.

Of these, one finds that Winslow and Gruenendahl (who often derives from the former) show the shape ெ which is attested for the fraction kāṇi ¹/₈₀ for the area kāṇi also. However, this is not seen in any other source and seems to be only a case of re-application of the fraction's symbol given the identical name. However, the area name mā is also that of a fraction, but that fraction's symbol ூ is not used for the area. Further, in the fractions, mā ¹/₂₀ is 4 times the kāṇi ¹/₈₀ but in area the kāṇi is 4 times the mā. It is hard to make out a clear interrelationship in this situation. All in all, we do not encode an area measure kāṇi.

Considering the highest area measure vēli, a sign for this is only given by Pope, Wickremasinghe and Kanita Nuul. The first two show merely the syllable ை (which is quite surprising considering that these two sources otherwise provide the most authentic and proper glyphs). Only Kanita Nuul shows the ligated glyph ை.

§4.8.2. Agriculture (see §3.4.2)

Among agricultural symbols, there is significant variation in the glyphs of the abbreviations for wet cultivation and dry cultivation. The final chosen glyphs are ை and ை respectively after Wickremasinghe and Pope's recent AES edition. However, Pope's older 1859 edition, Winslow (and Gruenendahl after him) list ை and ை. The latter especially is quite to be expected as a ligature of ூ (pu, for puncey) with the generic Tamil abbreviation mark ை. However, the former, though a ligature of Tamil na ை and na ை, is attested by Wickremasinghe for nāyakaṇ. While that attestation needs to be verified and is not proposed now (see §4.9), we have selected the glyph ை, which is seen only for nancey to avoid potential future conflicts.

In passing, we also note that Arden has used a couple of disjointed glyphs for these two signs. They may be disregarded for attestation.

Next, the sign alam is not per se related to agriculture but is grouped here in relation with kinds of land. It is translated by the sources as “salt pan”. The contemporary word for this is uppaḷam = uppu + aḷam where uppu = salt. Wickremasinghe shows the proper shape for this ை. Winslow has sought to approximate the shape below the ை by the letter LA ை, so: ை. Gruenendahl apparently followed him.

§4.8.3. Clerical and other abbreviations (see §3.4.3/4)

Two clerical characters, the traditional credit sign ௪ and number sign ௩ are discussed in §5.2 as there are existing characters ascribed to the same meaning. Three interrelated characters and one special character are separately discussed in the next two sections.

Of the remaining characters, most are straightforward and clearly identifiable as abbreviations from the corresponding Tamil words. A few cases are however to be noted:

The difference between *cirañcīvi* சீ and *cilvāṇam/cillarai* சீ is seen in Wickremasinghe and Pope. That between *pōka* பீ and *pillai* பீ is seen in Winslow and Gruenendahl. The former sources do not have *pōka* and the latter do not have *cilvāṇam*, but what they have, they clearly differentiate.

Wickremasinghe and Pope do not clearly differentiate *vacam* (“charge”) வசு – which is from the letters VA வ and CA ச – from the word-ending abbreviation *vum* வூ although they list them separately. Winslow and Gruenendahl do not list *vum*, but clearly derive *vacam* from VA and CA as would be expected.

The sign ரீ reads *rāja* (or *irāca*, lit. “royal”) is specific to one honorific phrase *rāja-rāja-śrī*, hence written as ரீ-ரீ-ஸ்ரீ . (Often it is prefixed by the word *mahā* மஹா .) This symbol is probably a ligature of the syllable *rā* in its old form: ரர , adding in the ubiquitous abbreviation mark ௨ to produce ரீ . It is often simplified and written as just ர (and often hyphenated, giving மஹா-ர-ர-ஸ்ரீ). However ர is also the shape of OBF1 TAMIL NUMBER ONE HUNDRED. This leads us to surmise that it was probably a later substitution for lack of printer’s repertoire, especially since only Wickremasinghe and Pope who provide the oldest and best attestations show the obviously older and original shape ரீ .

The original distinct shape is proposed for encoding. Those desiring the substitute shape may handle it with fonts or just use the existing character for ONE HUNDRED.

In passing one notes: Wickremasinghe and Pope both do not show the ஸ்ரீ in the Tamil text, but in the transliteration incorrectly write *śrī* as *stiri* (“woman”). Pope AES shows a small ௨ between the two ரீ -s. Wickremasinghe has it ligated with the first ரீ .




Two abbreviations, யூ for *yum* and வூ for *vum* are used in case of the word ending “-um” meaning “and”. With this, *malaiyum maṭuvum* would be written மலையூ மடுவூ .

A rare end-of-text mark ௨ is possibly related to the SINHALA KUNDDALIYA encoded at ODF4. Curiously, it also seems to have a distant echo in Newar (L2/12-003R §3.12 3).


§4.8.4. Mutal, mutaliya and vakaiyarā (see §3.4.4)


These three abbreviations are treated separately as they are interrelated.




The finalized shapes and meanings for encoding are:

mutal (“first”)  mutaliya (“and so on”)  vakaiyarā (“and such”) 

The difference between the first and second characters on one hand and the second and third on the other is to be noted. The glyphic difference is quite clear, but the meanings are to be carefully distinguished, especially between the latter two. The glyph of the second character and the meanings provided in the sources also need to be reviewed.

Regarding the attestations of the above three characters, there are many surprises with expected sources failing and unexpected ones stepping up. Wickremasinghe and Winslow list the two except mutaliya. Pope lists only vakaiyarā but not the other two. Gruenendahl list all three, and provides  as the glyph for mutaliya.

We also have an unexpected source for mutaliya in Beschi who has actually used this as part of running text. (Due to disuse, most of these attestations for characters are from reference texts.) Arden also lists mutaliya. Both Beschi and Arden do not list the other two symbols. Both have printed mutaliya as the syllable mi with an underline: .

Comparing the three shapes, one concludes that  of Beschi and Arden is merely an attempt to approximate the actual shape which should be formed by extending the stroke on the right side to go below and left. Gruenendahl’s shape on the other hand has this curved stroke right but has lost the distinctive shape of the central MA . From these indications we have reconstructed the shape of mutaliya as .

As for the meanings, Winslow and Gruenendahl after him list “from the first” for mutal. Wickremasinghe however has “first” which is more generic and hence taken for the character name. (Others do not attest mutal.)

For vakaiyarā, all four sources (the above three and Pope) read “belonging to”, but Wickremasinghe adds “et cætera” (sic). However, Gruenendahl, Beschi and Arden all translate mutaliya as “et cetera”. In fact, in native Tamil usage both vakaiyarā and mutaliya are different renderings of et cetera. mutaliya is used when the items are in a series with a recognizable order. vakaiyarā is used when the items just belong to a family or kind but no clear order can be defined. (We should however note that vakaiyarā has largely fallen out of common use and mutaliya is used in its place. So the distinction is not so clear nowadays.)

§4.8.5. K·KU (see §3.4.4)

The proposed 11FF4 TAMIL ABBREVIATION K·KU needs to be separately considered. While the Tamil script today does not employ any ligatures except K·SSA and SH·RII, the vagaries of fast writing has in the past produced some ligatures* which are to be seen in inscriptions and manuscripts. Wickremasinghe lists some of these ligatures on p 13:

க *kka* கூ *tta*
கூ or கூ *kku* கூ *ṭṭa*

Of these ligatures, K·KU is relatively the most common, probably since “ku” is the dative suffix in Tamil. Its ligated form, the first one to be precise, is still occasionally seen in handwritten Tamil text (sometimes even publicly visible and painted on walls), often with a pulli (for the vowelless initial KA) on top. However, the second shape shown here for K·KU has a different glyphic structure and a different usage pattern. It is hence proposed for encoding, although with another variant shape as representative glyph. (More below.)

Now Wickremasinghe himself separately shows similar shapes to the above in his list of Tamil signs. He labels them “varaikkum” in Tamil and “until” in English:

கூ or கூ = வரைக்கும் *varäikkum, until.*

The first shape here is essentially the same as the first shape shown above but with a longer tail that also turns slightly towards its end. The second shape is likewise also similar to the second one above, but without the loop’s overlap in the lower part of the glyph.

G U Pope shows a pair with the same labels (left, 1859; right, AES):

வரைக்கும் <i>until</i> கூ or கூ	வரைக்கும் <i>until</i> கூ or கூ
---	---

The second shape here has simplified the loop to a line, but this line now overlaps with the large curve encircling the center. (We note that second glyph from the later AES printing of Pope is slightly lacking, but this is probably only a random printing artifact.)

Here it is important to note that the first glyphs in all three pairs above, and the K·KA (other than K·KU) ligature shown by Wickremesinghe, has the core KA shape க retained as the first and core element of the ligature, with a short inverted cup forming the

* These ligatures need no support at the encoding level. Fonts may just provide them and other such attested written forms by glyph substitution of the C + VIRAMA + C sequence just as for K·SSA.

second component standing for the second KA in the underlying language content. This is just as in the Malayalam ligature K·KA கூ (which is prevalent even today). Thus we deem it appropriate to regard these first shapes written forms as mere ligatures to be handled by fonts by glyphic substitution. The second shapes for K·KU however have the inverted cup as the first glyphic feature and an approximation of the க as the second. By itself, this would neither be so serious a matter nor ground for distinct encoding, except for the fact that these second shapes have a distinct usage pattern which the first do not, as we shall show. Later scholars also have disregarded the first shape for a distinct symbol:

Winslow presents the same (second) shape as Pope:

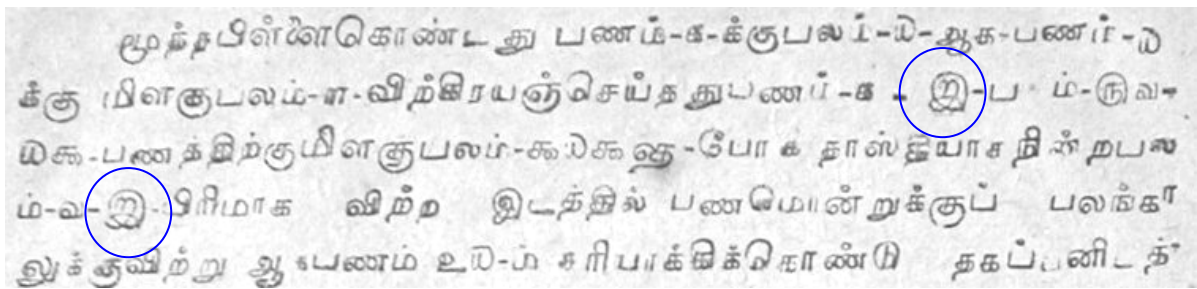
க்கு.....Until கு

... and of course Gruenendahl follows him, but erroneously spells it out as “kcu” (for “kku”).

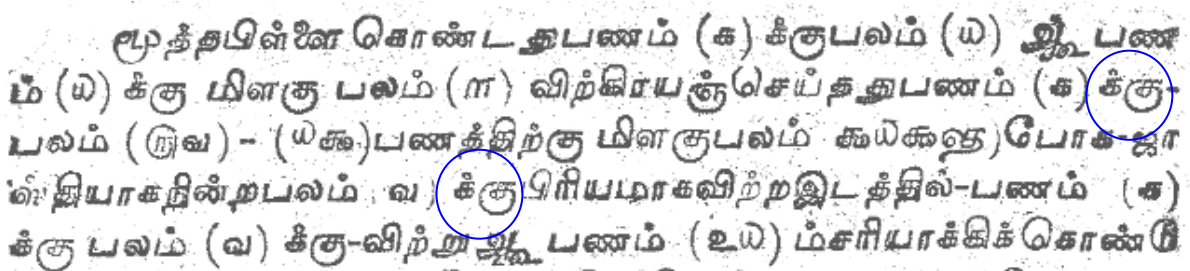
Ⓜ = க்சு until

Having said this much, we have to however note that it is unclear how these shapes have been labeled as varaikkum/“until” by the various authors. The only actual usage of these shapes seen so far have been in the original dative sense of “kku” only. Hence the proposed name is as the Tamil abbreviation for K·KU. (If usage for “until” may be confirmed later on, an annotation may well be added to the same character.)

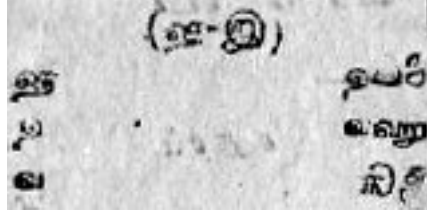
Now for the usage. Observe the following passage from Kanakkadigaram 1880:



... and the same passage (with some slight editorial changes) from Kanakkadigaram 1928:



Two relevant places have been circled. In the first instance, the text (in both editions) reads “paṇam **1kku** palam 5¼” and in the second instance “palam ¼**kku** piriyaṁāka virra iṭattil”. The older edition uses the glyph ௩ where the later edition (again, possibly due to loss of printer’s repertoire) writes out the full word ending kku: க்கு. Another example, from a text Neṭuṅkaṅakku Eṅ Cuvāṭi:

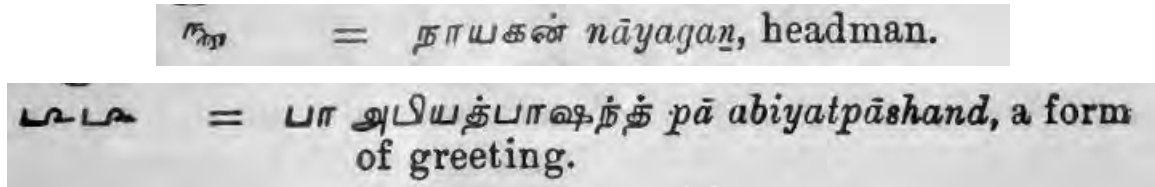


This reads: “ $3/4$ kku $3/4 = 1/2 + 1/20 + 1/80$ ” i.e. “ $3/4 \times 3/4 = 1/2 + 1/20 + 1/80$ ” and so on. (These are multiplication tables for fractions).

Thus it is clear that the character is indeed an abbreviation of the dative suffix kku used especially with numerals. It is proposed to be encoded for this purpose. The glyph chosen is ௩ based on Pope and Winslow, as the variant ௩ is confusable with I ௩.

§4.9. Characters not encoded pending further research

Wickremasinghe lists two symbols “nāyakaṅ” which means leader and “pā apiyatpāṣant” the meaning of which is not clear. These are not attested in any of the other sources:



I have noted in §4.8.2 that the glyph for nāyakaṅ is identical or highly similar to that shown by Winslow for naṅcey i.e. wet cultivation and hence the other shape was chosen for naṅcey. It is better to verify this shape for nāyakaṅ before encoding it. As for the other character, neither the meaning nor the usage is understood. Therefore further research is needed on the above two characters.

Kanita Nuul (p 298) shows symbols for three more weights: kaḷaṅcu, palam and eṭai.

10 பணவெடை - ௩	கழஞ்சு	5.2 கிராம்
15 கழஞ்சு - ௩	1 பலம்	78 கிராம்
20 பலம் - ௩	1 எடை	1 கிலோ 560 கிராம்

It also has a possible distinct attestation for the measure nāli/paṭi:



... which is attested in all other sources only as ஁ and hence not disunified from 0BF3 TAMIL DAY SIGN. One cannot discount the chance that the glyph ஁ for nāli/paṭi is only a simplification due to insufficient printer's repertoire.

As the book Kanita Nuul was based upon very many manuscripts as noted in its bibliography on p 275, it is likely that these and many more symbols are to be found in those manuscripts. Further research into those and other relevant manuscripts can produce the supplementary evidence needed to encode those characters later on.

If proper evidence for a distinct sign for the measure nāli/paṭi is found, it may be encoded in the place left reserved for it at 11FD9 between mūvuḷakku and kuruṇi.

§5. In relation to existing characters

§5.1. Disunification of characters with same shape

As noted in the previous sections, whenever possible we have chosen those written forms of characters which are not confusable with existing or other also-proposed characters. However, this is not possible when two characters are exclusively attested to be written the same way in contemporary texts. In these cases, the same written shape is disunified into two characters because the character properties, especially the GC, would be different. (This was already mentioned in §3.5.) The list of disunified characters follows:

Glyph	Existing character	Proposed character	
𑌒	0BAA LETTER PA	Lo	11FC8 FRACTION ONE-TWENTIETH No
𑌖	0BB5 LETTER VA	Lo	11FCF FRACTION ONE-QUARTER No
𑌛	0B99 LETTER NGA	Lo	11FDA SIGN KURUNI So
𑌔	0BA4 LETTER TA	Lo	11FDD SIGN TUUNI So
𑌗	0BB3 LETTER LLA	Lo	11FDE SIGN KALAM So
𑌚	-	-	{ 11FD1 FRACTION THREE-QUARTERS No 11FD6 SIGN UZHAKKU So

§5.2. Disunification of characters with same meaning

As previously mentioned in §4.8.3, two symbols are discussed here as they are identical in meaning to that currently documented for existing characters.

§5.2.1. Number sign vs balance sign

The sign ஈ is proposed for separate encoding as the TAMIL TRADITIONAL NUMBER SIGN. There is already a TAMIL NUMBER SIGN at 0BFA and its glyph as currently shown in the code chart is:

0BFA ஈ TAMIL NUMBER SIGN

However, this is the glyph provided for the balance sign in scholarly attestations, which contrast it with the number sign ஈ as currently proposed for separate encoding.

Wickremasinghe:

ஈ	=	நம்பர் <i>nambar</i> , number.
ஈ	=	நிலுவை <i>niluväi</i> , remainder or balance ; arrears.

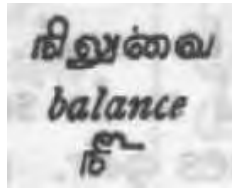
Winslow:

நம்பர்.....	A number.....	ஈ
நிலுவை.....	Balance	ஈ

Gruenendahl (who shows more than Winslow in this case):

ஈ & ஈ	=	நட்டபர் number
ஈ	=	நிலுவை balance

Pope (who does not list number but shows the current “number sign” for balance):



Arden (who does not list balance but shows an approximation of ஈ for number):

Number நம்பர் (English word) ஈ

It should be clear that scholars have distinguished the two characters.

One can readily see as how the number sign ஈ is derived from ந, the first letter of the word “number” written in Tamil viz நம்பர் nampar, whereas the balance sign நீ is likewise derived from நி, the first letter of the Tamil word for balance viz நிலுவை niluvai.

While Gruenendahl shows a நீ-like glyph ஈ for “number”, this is still however different from நீ and similar to ஈ in that it rests the abbreviation mark ஁ on the baseline on the right. However, ஈ more readily derives from ந (for நம்பர்) + ஁ than ஈ, and except Gruenendahl nobody else attests it for this sign. (The attestations shown in L2/01-375R which proposed நீ for the number sign were all from non-academic sources. Academic sources apparently were not consulted during that proposal.)

It certainly seems incongruous that number நம்பர் should be abbreviated as நீ. It may perhaps safely be deduced that this glyph for number is probably a later development (seen in the non-academic attestations) due to fine distinctions being forgotten.

Thus it is concluded that நீ was traditionally the balance sign and ஈ the number sign. We hence propose to encode ஈ by the name TRADITIONAL NUMBER SIGN and annotate the present character OBFA TAMIL NUMBER SIGN நீ for its original and later meanings.

§5.2.2. Credit sign vs income-set-aside sign

The sign ஁ is proposed for separate encoding as the TAMIL TRADITIONAL CREDIT SIGN. There is already a TAMIL CREDIT SIGN at OBFA and its glyph as currently shown in the code chart is:

0BF7 ஁ TAMIL CREDIT SIGN

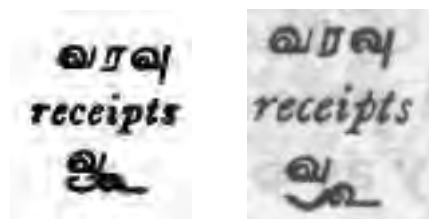
As will be mentioned in §5.3, this glyph should be corrected to ஁. (Apparently this correction has already been done. Michael Everson, email on Unicore dated 2012-May-02.)

However, the scholarly sources all list a different sign for the credit sign:

Wickremasinghe:

஁ = வரவு *varavu*, receipts.

Pope (left, 1859; right, AES):



Winslow:

வரவு....Receipts.... வு

Gruenendahl:

வவு = வு வு receipts

Clearly, Wickremasinghe and Pope 1859 have the original glyph, and the others approximated it with other available glyphs. Gruenendahl probably followed Winslow.

The point to note here is that all these academic sources derive the glyph for the varavu (credit) sign வவு from the first letter of the word viz VA வ . One also notes that the glyph for the parru (debit) sign பு (again see §5.3) is also derived from the first letter of the word viz PA ப . As such, it is quite surprising to see a glyph வு for the credit sign which instead derives from the unrelated letter E எ . (Again, the original proposal L2/01-375R only shows non-academic attestations for this எ -based glyph வு .)

While no contrastive description between வவு and வு is found in the scholarly sources obtained so far, as they nowhere even mention வு , feedback from people who have actually known or used these signs (various auditors, accountants and merchants of my acquaintance, personal communication) indicates that the VA-based sign வவு was indeed used for varavu = credit, whereas the E-based sign வு was used to mark another concept called eṭuppu , the first letter of which is indeed E எ . Apparently eṭuppu is income that is “ $\text{eṭuttu vaikkappaṭṭatu}$ ” i.e. “set aside” for unknown potential unexpected expenses. Later on for the total account to check out, this must be balanced by a proper debit entry (either towards expenses or towards cash deposited in the bank along with other income).

For some reason, apparently this “income set aside” sign வு was sometimes (later on?) used instead of the proper income i.e. credit sign வவு . While the attestation for வு in the previous proposal got reduced in quality when integrated into the PDF and hence I was not able to verify therefrom the usage of வு , I have obtained the actual scans, originating from Shri Chellappan of INFITT WG02 and kindly forwarded by Muthu Nedumaran, the then chairperson of INFITT WG02 which do show a shape like வு .

As such, the evidence for neither வவு nor வு can be rejected, and there is clear indication that these are semantically distinct. We hence propose to encode வவு by the

name TRADITIONAL CREDIT SIGN and annotate the present character 0BF7 TAMIL CREDIT SIGN ௪௩ for its original and later meanings.

§5.2.3. Rationale for proposed action plan

We initially considered changing the current glyph 0BFA TAMIL NUMBER SIGN from ௪௪ to ௪௩ and encode a separate balance sign character for ௪௪ . Likewise with changing 0BF7 TAMIL CREDIT SIGN to ௪௩ and encoding an “income-set-aside” character for ௪௩ .

This is however inadvisable as it is the written form that is encoded and not the meaning. Further, many Tamil Unicode fonts were produced post-4.0 (when these characters ௪௩ and ௪௪ were encoded) and probably based their glyphs for 0BF7 and 0BFA upon the chart. One had better not pull the carpet out from under them, as it were. While the glyph corrections already proposed in L2/12-106 §2 and mentioned in the next section are relatively minor, effectively moving written forms from one codepoint to another is inadvisable, especially from the viewpoint of font-makers and maintainers.

As such, the above action plan – of not changing the glyphs of the existing characters but annotating them for their actual usage, and adding new characters for the distinct written forms ௪௩ and ௪௩ – is found advisable.

§5.3. Glyph corrections for existing characters

The research for this proposal has indicated that two existing characters need their glyphs to be changed to reflect the actual attestation in scholarly sources. These have been documented in L2/12-106 §2. They are:

- 1) the change of the glyph of 0BF6 TAMIL DEBIT SIGN to ௪௪ where its chief component should clearly reflect its derivation from the Tamil PA ௪
- 2) the change of the glyph of 0BF7 TAMIL CREDIT SIGN to ௪௩ where its second component ௩ is the generic Tamil abbreviation marker

Attestation (from the same sources as here) has been provided in that document.

§5.4. Other matters concerning existing characters

We have already noted that the measure nāḷi/paṭi is not disunified from 0BF3 TAMIL DAY SIGN ௪ . That character must hence be annotated appropriately.

The same character om is also commonly used (see p 53) as a sign of auspiciousness, an invocation to the Hindu god Ganesha, the remover of obstacles. It is hence termed in Tamil “ pillaiyār sulī ”, literally “the whorl of Ganesha”, referring to the spiral in its shape.

Gruenendahl terms this “om”, presumably because it is used where the sacred syllable “om” would be used in texts printed in other scripts. Indeed, 0BD0 TAMIL OM om is mostly not used in such cases, and writing om to invoke Ganesha (an image of whom may be seen almost ubiquitously, even in a humble rural treeshade, in Tamil Nadu) is preferred for this. (It also helps that Ganesha is also termed the piraṇavakkaṭavuḷ , “the god of OM”.)

In line with characters of invocation found in other Indic scripts, such as the MEETEI MAYEK ANJI (AAF2), TIRHUTA ANJI (L2/11-175R), NEWAR SIDDHI and NEWAR SVASTI (L2/12-003R), BENGALI ANJI (L2/12-121) and DEVANAGARI SIDDHAM (L2/12-123), (which are all also associated with Ganesha and OM), this character may also be annotated as an invocation.

Other annotations were already mentioned in §5.2.

§6. Review of the GOTN standard L2/09-318

- 1) Additional characters not present in this standard (including some fractions) but newly encoded by this proposal should be added.
- 2) On pp 27, 28, 38 and 39 of the GOTN standard, mappings to Unicode are given. However, three missing mappings to existing Unicode characters viz. that of:

E108 ‘TAMIL FULL MOON SIGN’ to ○ 1F315 FULL MOON SYMBOL

E109 ‘TAMIL NEW MOON SIGN’ to ● 1F311 NEW MOON SYMBOL

E10A ‘TAMIL STAR SIGN’ to ★ 2605 BLACK STAR

... should be added. The remaining unmapped characters should map to characters newly encoded by this proposal.

- 3) P 43 intends to list the collation order for fractions. However, the fractions are in no recognizable meaningful order. It is better to have them in ascending order like the numerals as we have recommended later on in this document.
- 4) The glyph corrections described above in §5.3 need to be applied.

From Beschi p 123:

Sign.	Name.	Power.
வந்	முந்நீரை is the three hundred and twentieth part of the Integer,	$\frac{1}{320}$
ந	அரைக்காணி	$\frac{1}{160} = \frac{2}{320}$
ரி	காணி	$\frac{1}{80} = \frac{4}{320}$
சு	அரைமா	$\frac{1}{40} = \frac{8}{320}$
சூ	முக்காணி	$\frac{3}{80} = \frac{12}{320}$
ப	மா or ஒருமா	$\frac{1}{20} = \frac{16}{320}$
பரி	மாகாணி or வீசம்	$\frac{1}{16} = \frac{20}{320}$
சி	இரண்டிமா or இருமா	$\frac{1}{10} = \frac{32}{320}$
ஸ	அரைக்கால் or இரண்டிமாவரை	$\frac{1}{8} = \frac{40}{320}$
ந	மும்மா or முன்றுமா	$\frac{3}{20} = \frac{48}{320}$
சு	நாலுமா	$\frac{1}{5} = \frac{64}{320}$
வ	கால்	$\frac{1}{4} = \frac{80}{320}$
உ	அரை	$\frac{1}{2} = \frac{160}{320}$
ஞ	முக்கால்	$\frac{5}{4} = \frac{240}{320}$

க	1	ஒன்று
ந	2	முக்கால்
உ	3	அரை
வ	4	கால்
சி	5	நாலுமா
ஐ	6	அரைக்கால்
உ	7	இருமா
ஈ	8	மும்மாக்காணி
ப	9	மாக்காணி, வீசம்
நி	10	மும்மா
ப	11	ஒருமா
சு	12	அரைமா
சூ	13	முக்காணி
உ	14	காணி
லி	15	அரைக்காணி

1

கீழ்	உ	3	முத்திரி
கீ	ந	$\frac{2}{12.80}$	கீழ்முக்கால்
கீ	உ	$\frac{1}{12.80}$	கீழ் அரை
கீ	வ	$\frac{1}{12.80}$	கீழ் கால்
கீ	சி	$\frac{1}{12.80}$	கீழ் நாலுமா
கீ	ஈ	$\frac{2}{6.120}$	கீழ் மூன்றுவீசம்

From Shuddhananda Bharati p 204:

கீ	நி	$\frac{3}{6,400}$	கீழ் மும்மா
கீ	ஶு	$\frac{1}{2,500}$	கீழ் அரைக்கால்
கீ	உ	$\frac{1}{3,200}$	கீழ் இருமா
கீ	பந	$\frac{1}{5,120}$	கீழ் வீசம்
கீ	ப	$\frac{1}{6,400}$	கீழொருமா
கீ	சு	$\frac{3}{25,600}$	கீழ் முக்காணி
கீ	சு	$\frac{1}{12,800}$	கீழரைமா
கீ	ஸி	$\frac{1}{25,600}$	கீழ்க்காணி
கீ	நி	$\frac{1}{5,12,000}$	கீழ் அரைக்காணி
கீ	ஶு	$\frac{1}{1,02,400}$	கீழ் முத்திரி
	இம்மி	$\frac{1}{10,75,200}$	
	அதிசயம்	$\frac{1}{18,38,400}$	

From Burnell plate XXIII:

FRACTIONS								
$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{5}{16}$	$\frac{12}{16}$	$\frac{1}{20}$	$\frac{1}{30}$	$\frac{1}{320}$
പ ൧ ൧൧			വ	ര	ആ	പ	ഉ	൧൩

From Gruenendahl p 58:

ആ	൪	൧	൧൧ & ൧൧
$\frac{1}{320}$	$\frac{1}{160} = \frac{2}{320}$	$\frac{1}{80} = \frac{4}{320}$	$\frac{1}{40} = \frac{8}{320}$
൧൧൪	൧൧	൧	൧൧ & ൧൧
$\frac{1}{32} = \frac{8}{320} + \frac{2}{320}$	$\frac{3}{80} = \frac{8}{320} + \frac{4}{320}$	$\frac{1}{20} = \frac{16}{320}$	$\frac{1}{16} = \frac{16}{320} + \frac{4}{320}$
൧	൧൧ & ൧൧	൧൧	൧
$\frac{1}{10} = \frac{32}{320}$	$\frac{1}{8} = \frac{40}{320}$	$\frac{3}{20} = \frac{48}{320}$	$\frac{1}{5} = \frac{64}{320}$
൧	൧ & ൧	൧	
$\frac{1}{4} = \frac{80}{320}$	$\frac{1}{2} = \frac{160}{320}$	$\frac{3}{4} = \frac{240}{320}$	

From Caa Ganesan, p 16

தமிழ் எண்ணும், இலக்கமும். கல்வெட்டுக்களில் கண்டவண்ணம்
[சா. கணேசன்]

இந்திய எண்	A. D. 2. C.	A. D. 3. C.	A. D. 4. C.	A. D. 10. C.	A. D. 11. C.	A. D. 13. C.	A. D. 15. C.	A. D. 18. C.	A. D. 20. C.
கிரேக்க எண்	கி.மு 2. ப	கி.மு 3. ப	கி.மு 4. ப	கி.மு 10. ப	கி.மு 11. ப	கி.மு 13. ப	கி.மு 15. ப	கி.மு 18. ப	கி.மு 20. ப
$\frac{1}{16}$?	?	?	ஹ	ஹ	ஹ	ஹ	ஹ	ஹ
$\frac{1}{8}$?	?	?	ஹு	ஹு	ஹு	ஹு	ஹு	ஹு
$\frac{1}{4}$?	?	?	ஹ	ஹ	ஹ	ஹ	ஹ	ஹ
$\frac{1}{2}$?	?	?	ஹு	ஹு	ஹு	ஹு	ஹு	ஹு
$\frac{3}{4}$?	?	?	ஹு	ஹு	ஹு	ஹு	ஹு	ஹு

From Arden p 121:

Fractions

* 219. The following table gives the fractions most commonly used in Tamil; but, as already stated, Arabic figures are now generally used,

$\frac{1}{2}$	ஹு	அரை
$\frac{1}{4}$	ஹ	கால்
$\frac{1}{8}$ (i.e. $\frac{1}{2}$ of $\frac{1}{4}$)	ஹு	அரைக்கால்
$\frac{3}{4}$ (i.e. 3 times $\frac{1}{4}$)	ஹு	முக்கால்

From Subramanian pp 81, 82:

2	4	9	5	ஹு	ஹ	ஹ	ஹு
2	4	9	5	$\frac{3}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$

களிலேதான் அதிகமாக உள்ளன. படம் 7-இல், கி. பி. 1204-ஆம் ஆண்டில் வழங்கிவந்த $\frac{3}{4}$, $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, ஆகியவற்றின் உருவங்கள் காட்டப் பட்டுள்ளன. இவற்றைத் தவிர, காணி முந்திரி போல்வன வற்றிற்கும், வடிவங்கள் காணப்படுகின்றன.

From Beythan p 147:

2. Einige Bruchzeichen

Wert	Zeichen	Name	Wörtl.
$\frac{1}{20}$ *)	ஞ	முந்திரி	<i>muntiri</i>
$\frac{1}{80} = \frac{1}{20}$	ற	அரைகாணி	<i>araikāṇi</i>
$\frac{1}{80} = \frac{1}{20}$	ஐ	காணி	<i>kāṇi</i>
$\frac{1}{80} = \frac{1}{20}$	ச	அரைமா	<i>araimā</i>
$\frac{3}{80} = \frac{1}{20}$	சூ	முக்காணி	<i>mukkāṇi</i>
$\frac{1}{80} = \frac{1}{20}$	ப	மா	<i>mā</i>
$\frac{1}{80} = \frac{1}{20}$	பழ	மாகாணி (வீசம்)	<i>mākāṇi (vīcam)</i>
$\frac{1}{80} = \frac{1}{20}$	உ	இருமா	<i>irumā</i>
$\frac{1}{4} = \frac{1}{20}$	ஞ	அரைக்கால்	<i>araikkāl</i>
$\frac{1}{4} = \frac{1}{20}$	சி	நாலுமா	<i>nālumā</i>
$\frac{1}{4} = \frac{1}{20}$	வ	கால்	<i>kāl</i>
$\frac{1}{4} = \frac{1}{20}$	உ	அரை	<i>arai</i>
$\frac{3}{4} = \frac{1}{20}$	த	முக்கால்	<i>mukkāl</i>

From Winslow p 976:

அண்ணாAn anna	அண்ணா
அரைமா $\frac{1}{40}$	ச
அளம்Salt pan.....	அல
ஆகTotal.....	ஆக
இரண்டுமா $\frac{1}{80}$	உ
மா $\frac{1}{20}$	ப
மாகாணி $\frac{1}{80}$	பழ
மாசம்A month.....	மீ
முக்காணி $\frac{3}{80}$	சூ
முதல்From the first	மே
முந்திரி $\frac{1}{20}$	வத

தமிழ்ப் பின்ன இலக்க விளக்கம்

வ	முந்திரி	$\frac{1}{320}$	ஐ	இரண்டுமா	$\frac{1}{16}$
ஈ	அரைக்காணி	$\frac{1}{160}$	ஈ	மூன்றுமா	$\frac{3}{20}$
உ	காணி	$\frac{1}{80}$	ஊ	நான்கு	$\frac{1}{5}$
ஈ	அரைமா	$\frac{1}{40}$	ஊ	கால்	$\frac{1}{4}$
ஊ	முக்காணி	$\frac{3}{80}$	உ	அரை (இ)	$\frac{1}{2}$
ஊ	ஒருமா	$\frac{1}{20}$	ஊ	முக்கால் (ஈ)	$\frac{3}{4}$
			ஊ	ஒன்று (முழு எண்)	1

நாணயப்பகுதி பின்ன எண்வகை

ஊ	கால் வீசம் (உவ)	$\frac{1}{64}$	ஈ	மூன்று வீசம் (ஈஊ)	$\frac{3}{16}$
ஈ	அரை ,, (ஈஈ)	$\frac{1}{32}$	ஊ	கால்	$\frac{1}{4}$
ஊ	முக்கால் ,, (ஊஈவ)	$\frac{3}{64}$	உ	அரை	$\frac{1}{2}$
ஊ	வீசம் (ஈ)	$\frac{1}{16}$	ஊ	முக்கால் (ஈ)	$\frac{3}{4}$
ஊ	அரைக்கால் (ஈஈ)	$\frac{1}{8}$	ஊ	ஒன்று (முழு எண்)	1

பின்ன இலக்க அட்டவணை

$\frac{1}{20}$	ஒன்றின் பாகம்	=	முந்திரி	$\frac{1}{320}$	வ
2	முந்திரி	=	அரைக்காணி	$\frac{1}{160}$	ஈ
2	அரைக்காணி	=	காணி	$\frac{1}{80}$	உ
2	அரைக்காணி	=	அரைமா	$\frac{1}{40}$	ஈ
3	காணி	=	முக்காணி	$\frac{3}{80}$	ஊ

From Kanakkadigaram 1958 p 14:

2	அரைமா	=	மா	$\frac{1}{20}$	ட
2	ஒருமா	=	இரண்டுமா	$\frac{1}{10}$	ஊ
3	”	=	மூணுமா	$\frac{3}{20}$	ஈ
2	இரண்டுமா	=	நாலுமா	$\frac{1}{5}$	சு
5	ஒருமா	=	கால்	$\frac{1}{4}$	வ
10	”	=	அரை	$\frac{1}{2}$	உ
15	”	=	முக்கால்	$\frac{3}{4}$	து
20	”	=	ஒன்று	1	க
நாணயப்பகுதி பின்ன இலக்க அட்டவணை					
5	முந்திரி	=	கால் வீசம்		ஊ
2	கால் வீசம்	=	அரை வீசம்		சூ
3	கால் ”	=	முக்கால் வீசம்		கூ
2	அரை ”	=	வீசம்	(ஸ)	பூ
2	வீசம்	=	அரைக்கால்		ஹ
3	”	=	மூணு வீசம்	(ஈ)	ஹீ
2	அரைக்கால்	=	கால்		வ
2	கால்	=	அரை	(இ)	உ
3	கால்	=	முக்கால்	(ஈ)	து
4	” அல்லது	}	ஒன்று (முழு எண்)		க
2	அரை				

கீழ்வாய் இலக்கம்

ன	(ன)	பு	வ	சு	பு	(பு)	வ
உ	வ	பு	பு	சு	கீ
வ	ந	சு	ந	சு	கீ
ந	வ	சு	(வ)	ந	பு	கீ
வ	பு	சு	வ	ந	கீ
பு	சு	ந	கீ	வ	கீ
சு	சு	வ	கீ	ந	(சு)	கீ
சு	பு	வ	கீ	ந	கீ
பு	ந	ந	(ந)	சு	பு	கீ
ந	வ	வ	பு	ந	கீ
வ	கீ	பு	ந	வ	கீ
பு	(பு)	வ	சு	ந	சு	(சு)	கீ
வ	வ	சு	வ	சு	கீ
ந	பு	பு	கீ	பு	கீ
வ	பு	ந	கீ	ந	கீ
பு	சு	வ	கீ	ந	கீ
சு	பு	வ	கீ	ந	கீ
சு	பு	வ	(வ)	பு	பு	(பு)	கீ
பு	பு	பு	ந	ந	கீ
ந	ந	சு	வ	வ	கீ
வ	கீ	சு	வ	ந	(ந)	கீ
ந	(ந)	பு	பு	கீ	வ	கீ
வ	சு	வ	கீ	வ	(வ)	கீ

பின்ன இலக்கம்

‘பின்ன இலக்கம்’ என்பது: ஒரு முழு எண்ணை பாகித்து (பின்னப்படுத்தி) உபயோகிப்பதற்கு வழங்கும் பெயராகும். தமிழில் ஒரு முழு எண்ணை நூஉய பாகங்க ளாகப் பகுத்துவரும் ஒரு பாகத்திற்கு (ஊ) முந்திரி என்று பெயர். முந்திரியும், அது கூடுதலாகும் தொகை யின் மதிப்பும் கீழே தரப்பட்டிருக்கின்றது.

தமிழ் பின்ன இலக்கம் உச்சரிக்கும் முறை

*ஊ முந்திரி	$\frac{1}{320}$	சய அரைமா	$\frac{1}{40}$
ர அரைக்காணி	$\frac{1}{160}$	சூ முக்காணி	$\frac{3}{80}$
ரூ காணி	$\frac{1}{80}$	ப ஒருமா	$\frac{1}{20}$
ஊ இரண்டுமா	$\frac{1}{10}$	உ அரை	$\frac{1}{2}$
ரு மூன்றுமா	$\frac{3}{20}$	*ஊ முக்கால்	$\frac{3}{4}$
கூ நாலும்மா	$\frac{1}{5}$	க ஒன்று (முழு எண்)	
வ கால்	$\frac{1}{4}$		

நாணயப் பகுதி பின்ன எண் வகை

ஊ கால் வீசம் (ரூஊ)	$\frac{1}{64}$	ரு- மூணு வீசம் (ருசூ)	$\frac{3}{16}$
சூ அரை ,, (சயர)	$\frac{1}{32}$	வ கால்	$\frac{1}{4}$
சூஉ முக்கால் ,, (சூரஊ)	$\frac{3}{64}$	உ அரை	$\frac{1}{2}$
பூ வீசம் (பரூ)	$\frac{1}{16}$	ஊ முக்கால்	$\frac{3}{4}$
ஊ அரைக்கால் (ஊசூ)	$\frac{1}{8}$	க ஒன்று (முழு எண்)	

From Wickremasinghe p 41:

$\frac{1}{80}$	ரூ	காணி <i>kāṇi</i>	$\frac{3}{20}$	ரு	மூன்றுமா <i>mūṇru-mā</i>
$\frac{3}{80}$	சூ	முக்காணி <i>mukkāṇi</i>	$\frac{4}{80}$	கூ	{ நாலுமா <i>nālu-mā</i> , or
$\frac{1}{20}$	ப	மா <i>mā</i> ; $\frac{1}{40}$ சய <i>arūi-mā</i>			i.e. $\frac{1}{5}$ { நான்குமா <i>nāngu-mā</i>
$\frac{1}{16}$	பரூ	{ மாகாணி <i>mā-gāṇi</i> , or	$\frac{1}{8}$	ஊ	அரைக்கால் <i>arūik-kāl</i>
		{ வீசம் <i>vīcam</i>	$\frac{1}{4}$	வ	கால் <i>kāl</i>
$\frac{2}{20}$	ஊ	{ இரண்டுமா <i>iraṇḍu-</i>	$\frac{1}{2}$	உ	அரை <i>arūi</i>
OR $\frac{1}{10}$		{ <i>mā</i> , or இருமா <i>iru-mā</i>	$\frac{3}{4}$	ஓ	முக்கால் <i>mukkāl</i>

§7.2. Tamil Symbols

(Some attestations for symbols were already provided along with the fractions in §7.1.)

From Kanakkadigaram 1958 p 21:

சாதாரணமாய்த் தமிழ் நாட்டில் தோன்றுதோட்டு வழங்கிவரும் அளவை		
360 ரெல்	= 1 செவிடு	சூ
5 செவிடு	= 1 ஆழாக்கு	ஶ
2 ஆழாக்கு	= 1 உழக்கு	சூ
2 உழக்கு	= 1 உரி	வரி
3 உழக்கு	= 1 மூவுழக்கு	நூ
2 உரி	= 1 நாழி	உ
8 நாழி (படி)	= 1 குறுணி (மரக்கால்)	ஈ
2 குறுணி	= 1 பதக்கு	ஶசூ
3 குறுணி	= 1 முக்குறுணி	ஈ
2 பதக்கு (4 குறுணி)	= 1 தூணி	த
2 தூணி	= 8 குறுணி	ஶசூ
3 தூணி	= 1 கலம்	சூ, ஶ
5 மரக்கால்	= 1 பறை	
80 பறை அல்லது 400 மரக்கால்	} = 1 கரிசை	

From Beschi pp 129 and 130:

கூஉ	-	ஒருச்சுவடு	-	-	-	-	1	Suvadu.
உஊஉ	-	இருச்சுவடு	-	-	-	-	2	
நஊஉ	-	மூச்சுவடு	-	-	-	-	3	
சஊஉ	-	நாற்சுவடு	-	-	-	-	4	
ஹ	-	ஆழாக்கு	-	-	-	-	5	Suvadu 1 Alhāccu.
ஊ	-	உழக்கு	-	-	-	-	2	
ஊஹ	-	உழக்காழாக்கு	-	-	-	-	3	
வரி	-	உரி	-	-	-	-	4	
வரிஹ	-	உரியாழாக்கு	-	-	-	-	5	
நு	-	மூவுழக்கு	-	-	-	-	6	
நுஹ	-	மூவுழக்காழாக்கு	-	-	-	-	7	
உ	-	நாழி	-	-	-	-	8	Alhāccu 1 Nālhi.
உஉ	-	இருநாழி	-	-	-	-	2	
நஉ	-	முந்நாழி	-	-	-	-	3	
சஉ	-	நானாழி	-	-	-	-	4	
ஊஉ	-	ஐந்நாழி	-	-	-	-	5	Nālhi.
சுஉ	-	அறுநாழி	-	-	-	-	6	
எஉ	-	எழுநாழி	-	-	-	-	7	
ஊ	-	குறுணி	-	-	-	-	8	Nālhi 1 Curuni.
உஊ	-	பதக்கு	-	-	-	-	2	
ஊஊ	-	மூக்குறுணி	-	-	-	-	3	
த	-	தூணி	-	-	-	-	4	
தஊ	-	ஐங்குறுணி	-	-	-	-	5	
தஊஊ	-	தூணிப்பதக்கு	-	-	-	-	6	
தஊஊ	-	எழுகுறுணி	-	-	-	-	7	
உத	-	இருதூணி	-	-	-	-	8	
உதஊ	-	இருதூணிக்குறுணி	-	-	-	-	9	
உதஊஊ	-	இருதூணிப்பதக்கு	-	-	-	-	10	
உதஊஊ	-	இருதூணிமூக்குறுணி	-	-	-	-	11	
சுஊ	-	கலம்	-	-	-	-	12	Curuni 1 Calam.

From Cintamani En Suvadi pp 30, 31:

நெல்லிலக்க வாய்பாடு		
360 நெல்	1 சவடு	ஆ
5 சவடு	1 ஆழாக்கு	ஹ
2 ஆழாக்கு	1 உழக்கு	ஞ
2 உழக்கு (அல்லது)	} 1 உரி	வரி
4 ஆழாக்கு		
*3 உழக்கு	1 மூ வெழாக்கு	நு

* உழக்கு, ஆழாக்கு என்பதைச் சரியாய்க் கவனித்து உச்சரிக்கவேண்டும்.

2 உரி (அல்லது)	} 1 நாழி (படி)	உ
8 ஆழாக்கு		
8உ (அல்லது)	} 1 மரக்கால் (குருணி)	ஈ
8 படி		
2 குருணி	1 பதக்கு	*ப (வது)
3 குருணி	1 முக்குருணி	நி
4 ,,	1 தூணி	த
2 தூணி	1 இரு தூணி	வது
3 தூணி (அல்லது)	} 1 கலம்	சுள (ள)
12 மரக்கால்		
21 மரக்கால்		

Abbreviations & Contractions

௨	=	om̐ (cf. below)
Ⓜ	=	க்சு until
ஞ	=	சிரஞ்சீவி long-lived
ஸ் & ஸ்ரீ	=	சிரீ śrī
பீ	=	பிள்ளை <i>pillai</i> , a title of rank
ம-ந-ந-ஸ்ரீ	=	மகாராஜ ராஜசிரீ <i>mahārājarājaśrī</i> , abbr. "M. R. Ry."
ஸ்ரீ	=	முதல் from the first
ஶ	=	வகையரா belonging to
ஷ	=	மேற்படி afore-mentioned
ஶ	=	முதலியவை "etc."
- <i>Accounting</i>		
ஶ	=	ஆக total
ந & ந்	=	நட்பார் number
நீ	=	நிலுவை balance

From Gruenendahl p 53:

TAMIL

53

பு = போக what is spent

ஶ = வசம் charge

வக = வரவு receipts

- *Agriculture etc.*

அல = அளம் salt pan

ஐ & று = நெல்லு paddy

நு = நன்செய் wet cultivation

நீ = நிலம் ground

பு = புன்செய் dry cultivation

- *Anglo-Indian Currency*

க = காசு a Cash

பு = புண்டம் a Fanam = 80 Cash

ஶ = பத்தன் ten Fanams

வக = வராகன் a Pagoda = 42 Fanams

பு = பைசா a Paise or Pie

அ = அண்ணா an Anna

ரூ & ரூ = ரூபாய் a *Rupee* = 16 *Annas*

- *Calendar*

வரு = வருஷம் year

மீ = மீசம் month

உ & உ = தேதி date

நா = நாள்து current

- *Weights & Measures*

பா = பாரம் a *Candy* (an "Indian ton")

ம = குறுணி one *Marcals* (measure of capacity)

உ = பதக்கு two *Marcals*

த = ஐங்குறுணி five *Marcals*

உ = இருதூணி eight *Marcals*

க = கலம் a *Kalam* = 12 *Marcals*

க = காணி a *Cawney* (measure of land)

க = கசம் a yard

க = குடி a square yard

From Winslow p 976:

மேற்படி.....	Above said.	மே
ரூபாய்.....	A rupee.....	ரூ
வகையரா	Belonging to.....	வ
வசம்.....	Charge.....	வ
வரவு.....	Receipts.....	வ
வராகன்.....	A pagoda.	வ
வருஷம்.....	A year.	வ

நெல்லு	Paddy.....	ந
பணம்.....	Fanam.....	ப
பதக்கு.....	Two marcals.....	ப
பாரம்.....	A candy.....	ப
பின்னே.....	A title of rank.....	ப
புன்செய்.....	Dry cultivation.....	ப
பைசா.....	A pie.....	ப
பொன்	Ten fanams.....	ப
போக.....	What is spent.....	ப

From Winslow p 976 (continued):

இருதூணி.....	Eight marcals.....	வது
ஐங்குறுணி.....	Five marcals.....	தங்
கலம்.....	A kalam.....	கூர்
காசு.....	A cash.....	கூ
காணி.....	A cawney.....	சி
குழி.....	A square yard.....	கூர்
குறுணி.....	One marcal.....	ங்
கெசம்.....	A yard.....	கூர்
க்கு.....	Until	மு
சொஞ்சிவி.....	Long lived.....	கூர்
தேதி.....	Date	உ
நஞ்செய்.....	Wet cultivation.....	கூர்
நம்பர்.....	A number.....	ங்
நாளது.....	Current.....	நூ
நிலம்.....	Ground.....	நூ
நிலுவை.....	Balance	நூ

APPENDIX I

ABBREVIATIONS

The following abbreviations are in common use :

English	Tamil word	Abbreviation
A Rupee	ரூபாய்	ரூ or ரூ ²
An anna ($\frac{1}{16}$ of a rupee)	அணா	அ
A pie ($\frac{1}{32}$ of an anna)	பை	பை
A year	வருஷம்	வரு
A month	மாசம்	மீ
Date (i.e. Day of the month)	தேதி	உ
Et cetera	முதலியவை	மி
Aforesaid	மேற்படி	செடி
Fortunate, Holy, Mr.	சிரீ	ஸ்ரீ
Honorable Mr. or M. R. Ry.	} மகாராஜ ராஜஸ்ரீ ம-ரா-ரா-ஸ்ரீ	
Current		நாளது
Number	நம்பர் (English word)	கூர்

The following and many others occur occasionally in documents written in old-fashioned style.

A <i>Kalam</i> measure (wet or dry)	} கலம்	சுரா
A weight of 500 lb.		பாரம்
A yard	கஜம்	கூ
A square yard	குழி	கூ
Wet cultivation	நன்செய்	நூ
Dry cultivation	புன்செய்	பூ
A Rupee	ரூபாய்	ரூ
Long-lived (a polite title to a junior)	} சிரஞ்சீவி	சூ

From Pope 1859 p 18:

VIII. The following abbreviations and signs are in use:

நன்செய் wet cultivation நீர்	புன்செய் dry cultivation பூ	வேலி a land measure வி	நாளது current நூ
மாசம் month மீ	தேதி day உ	வருஷம் year வூ or நூ	நிலம் ground நீ
நிலுவை balance நீ	பற்று received பூ	வரவு receipts வூ	வசம் charge வூ
மேற்படி aforesaid அ	பணம் money பு	காசு cash கூ	வராகன் a pagoda வூ
ரூபாய் rupee ரூ	பொன் 10 fanams பூ	கஜம் a yard கூ	வரைக்கும் until கூ or நூ
குழி a sq. yard கூ	ஆக total ஆ	அணா an anna அணா	பைசா a pie பூ
வகையரா belonging to வூ	சிலவானம் more or less சூ	நெல்லு paddy நூ	கலம் a measure கூ
சிரஞ்சீவி (long lived) a polite title for a junior சீ	இராச இராசஸ்திகி most honorable, a title நீ நீ	வும், யும் and (+ வ் or ய்) வூ, யூ	

From Pope AES:

14

INTRODUCTION

viii. The following abbreviations and signs are in use:—

நன்செய் <i>wet cultivation</i>	புன்செய் <i>dry cultivation</i>	வேலி <i>a land measure</i>	நாளது <i>current</i>
நூ	பறு	லி	நூ
மாசம் <i>month</i>	தேதி <i>day</i>	வருஷம் <i>year</i>	நிலம் <i>ground</i>
மீ	உ	வூ or யூ	நீ
நிலுவை <i>balance</i>	பற்று <i>received</i>	வரவு <i>receipts</i>	வசம் <i>charge</i>
நீ	பு	வூ	வூ
மேற்படி <i>aforsaid</i>	பணம் <i>money</i>	காசு <i>cash</i>	வராகன் <i>a pagoda</i>
சூடி	பு	கூ	வூ
ரூபாய் <i>rupee</i>	பொன் <i>10 fanamis</i>	கஜம் <i>a yard</i>	வரைக்கும் <i>until</i>
ரூ or தூ	பூ	கூ	கூ or நூ
குழி <i>a square yard</i>	ஆக <i>total</i>	அணா <i>an anna</i>	பைசா <i>a pie</i>
கூ	அகூ	அணூ	டபை
வகையரா <i>belonging to</i>	சில்வானம் <i>more or less</i>	நெல்லு <i>paddy</i>	கலம் <i>a measure</i>
கூ	கூ	பூ	கூ
கிரஞ்சிவி <i>a polite title for a junior</i>	இராச இராசஸ்திரி <i>most honourable, a title</i>		அம் யும் வூ யூ
கூ	ராசா (for <i>rāja-ṣrī</i>)		

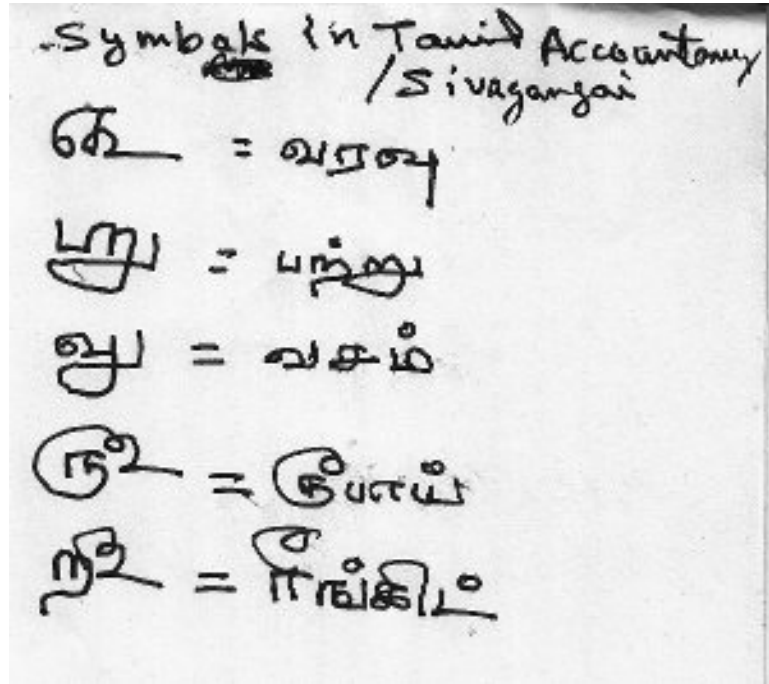
அ or ஆ	=	அணு <i>aṇā</i> , an anna.
ஆ	=	அளம் <i>aḷam</i> , a salt-pan.
ஆக	=	ஆக <i>āga</i> , total.
இராச ராசஸ்திரி	=	இராச ராசஸ்திரி <i>irāja rājastiri</i> , most honourable (a title).
கா	=	கலம் <i>kalam</i> , a measure of grain.
காடி	=	கஜம் <i>kajam</i> , a yard.
காசு	=	காசு <i>kāṣu</i> , cash.
காணி	=	காணி <i>kāṇi</i> , the fraction $\frac{1}{8}$.
குடி	=	குடி <i>kuṇḍi</i> , a square (yard).
சீ	=	சிரஞ்சீவ <i>ṣirañjīva</i> , "live long," a polite title used in addressing a junior in rank.
சீவா	=	சில்வாணம் <i>ṣilvāṇam</i> , and odd.
தேதி	=	தேதி <i>tēdi</i> , day of the month.
நம்பர்	=	நம்பர் <i>nambar</i> , number.
நன்செய்	=	நன்செய் <i>nanṣey</i> , wet cultivation.
நாயகன்	=	நாயகன் <i>nāyagan</i> , headman.
நாளது	=	நாளது <i>nāḷadu</i> , current.
நிலம்	=	நிலம் <i>nilam</i> , ground.
நிலுவை	=	நிலுவை <i>niluvāi</i> , remainder or balance; arrears.
நெல்லு	=	நெல்லு <i>nellu</i> , paddy.
பணம்	=	பணம் <i>paṇam</i> , money.
பற்று	=	பற்று <i>pattu</i> , received.
பாரம்	=	பாரம் <i>pāram</i> , weight of 500 lbs.
பா அபியத்பாஷந்த்	=	பா அபியத்பாஷந்த் <i>pā abiyatpāshand</i> , a form of greeting.
பிள்ளை	=	பிள்ளை <i>pillāi</i> , caste title.

From Wickremasinghe pp 14:

புறு	=	புன்செய் <i>punṣey</i> , dry cultivation.
பை	=	பைசா <i>paicā</i> , a pie.
பெ	=	பொன் <i>pon</i> , a gold coin equal to ten fanams.
மீ	=	மாசம் <i>mācam</i> , a month.
மு	=	முதல் <i>mudal</i> , first.
மே	=	மேற்படி <i>mētpadi</i> , aforesaid, as above.
யு	=	யும் <i>yum</i> , and.
ரூ or ரூ	=	ரூபாய் <i>rūbāy</i> , rupee.
வகை	=	வகையரா <i>vagāiyarā</i> , belonging to; <i>et cætera</i> .
வச	=	வசம் <i>vaṣam</i> , charge, control.
வரவு	=	வரவு <i>varavu</i> , receipts.
வரா	=	வராகன் <i>varāgan</i> , a pagoda coin.
வரு or வரு	=	வருஷம் <i>varuṣham</i> , year.
வரை or வரை	=	வரைக்கும் <i>varāikkum</i> , until.
வம்	=	வம் <i>vum</i> , and.
வே	=	வேலி <i>vēli</i> , a land measure about 5 acres.

From <http://jaybeesnotebook.blogspot.in/2012/02/tamil-accountancy-symbols-1.html>

retrieved 2012-Jul-07



From Kanita Nuul p 296:

12 அங்குலம்	- 1 அடி	300 மி.மீ; 30 செ.மீட்டர்
3 அடி (36 அங்குலம்)	- 1 கெஜம்	90 செ.மீ
144 சதுர அடி (16 சதுரகெஜம்)	- 1 குழி - கு	13 சதுர மீட்டர்
100 குழி (1600 சதுரகெஜம்)	- 1 மா	1,300 சதுர மீட்டர்
20 மா (32,000 சதுரகெஜம்)	- 1 வேலி லு	26,000 சதுர மீட்டர்

Specific usage attestations for individual symbols

For 0BF3 TAMIL DAY SIGN used as “pillaiyār sulī”, an invocation to the god Ganesha:
(from Kanakkadigaram 1928 inner front cover)

உ
சிவமயம்.
கணக்கதிகாரம்.

For TAMIL SIGN MUDALIYA (from Beschi p 122):

கய	-	-	முப்பது	-	-	-	-	-	30
கயக	-	-	முப்பத்தொன்று	யி	-	-	-	-	31 &c.
சய	-	-	நாற்பது	-	-	-	-	-	40
சயஉ	-	-	நாற்பத்திரண்டு	யி	-	-	-	-	42 &c.
டுய	-	-	ஐம்பது	-	-	-	-	-	50
டுயக	-	-	ஐம்பத்துமூன்று	யி	-	-	-	-	53 &c.
சுய	-	-	அறுபது	-	-	-	-	-	60
சுயச	-	-	அறுபத்துநான்கு	யி	-	-	-	-	64 &c.

For TAMIL SIGN RAAJA substituted by the glyph akin to Tamil 100 ற:

ம - ற - ற - றீ,
V. K. சுந்தரேஸ்வர முதலியாரவர்கள்

For TAMIL CURRENT SIGN:

ஆவது நம்பர்சன்னதுசாதராசகது.
பமோஜிப்டக்கும் பித்விநுயீ (கூஉ) சாயங்காலமஷ்கிராமத்துகருப்போய்
அதைசுற்றிப்பார்த்தபோது சாகுபடிகருலாயக்காகவும தரிசாகவுமஇருக்கிற
நிலங்களினவிஸ்தாரத்தைககண்டுகொண்டு முதலாவது ஷ்கிராமத்துககணக
கணைவரவழைத்து சாகுபடிககணகருகுகொண்டிவரசசொல்லிததாக்கீதுபண

For TAMIL SIGN KUZHI:

க-வது. இந்த டிஸ்திரிகட்டு முன்சிப் இலாகாவில் இருக்கிற நிலங்கள்
உச அடிகோலால், கோல் க-க்கு ஒரு குட யாக ௧௦௦ குட கரு காணி

For TAMIL TOTAL SIGN and TAMIL ABBREVIATION KKV:

ஆ மொத்தம் உதாரண சாத்திரங்களில் உரு-இ
தனிப்பெயராய் வந்த சாத்திரங்கள் கசுரு.
* கோட்டிற்கு மேலுள்ள பெயர்கள் கூ-சூ, முதல் பன்விர

For TAMIL PUNCTUATION END OF TEXT:

இததால் அரியபயண னுகிரதாவது ன

சு அளவு தவிர தாவிவன்கெ சிலவேண்டி தவிரியபயண விரிந்து தவிரிய
ரரு தவிரியபயண விரிந்து தவிரிய

The earliest available plakkaat printed in Tamil dated 22/09/1742.

§8. References

1. **Arden:** A Progressive Grammar of the Tamil Language, 5th Edition, A H Arden, revised by A C Clayton, Christian Literature Society, Madras, 1976
2. **Beythan:** Praktische Grammatik der Tamilsprache, Hermann Beythan, Leipzig, 1943
3. **Beschi:** A Grammar of the Common Dialect of the Tamul Language called Kotuntamil, Constantius Joseph Beschi, Tr: George William Mahon, Madras, 1848
4. **Burnell:** Elements of South Indian Palaeography, Arthur Burnell, London, 1878, <http://www.archive.org/details/elementssouthin05burngoog>
5. **Caa Ganesan:** Handbook issued at the 2nd International Tamil Research Conference, Ed: Caa Ganesan, Madras, 1968
6. **Cintamani En Suvadi:** Cintāmaṇi Eṇ Cuvaṭi, S Kalyāṇacuntaram Piḷḷai, C Kumarasamy Naidu & Sons, Chennai, 1950
7. **Ezhuttadigaram:** Tolkāppiyam Eḷuttatikāram, Tirumakaḷ Aḷuttakam, Cuṇṇāgam, Śrī Lankā, 1937, http://noolaham.org/wiki/index.php/தொல்காப்பியம்_எழுத்ததிகாரம்
8. **Gruenendahl:** South Indian Scripts in Sanskrit Manuscripts and Prints, ISBN 3-447-04504-3, Reinhold Gruenendahl, Wiesbaden, Germany, 2001
9. **Kanakkadigaram 1880:** Koṟukkaiyūr Kāri Nāyaṇār, Ed: Ārumuka Mutaliyār, Pub: Tiruvorriyūr Paracurāma Mutaliyār, Parappirama Muttirākṣara Cālai, 1880
10. **Kanakkadigaram 1928:** Koṟukkaiyūr Kāri Nāyaṇār, Pub: Nēlaṭūr Cuppiramaṇi Aiyar, Vittiyaṅvilāsa Muttirākṣara Cālai, Pirapava Varuṭam Paṅkuṇi Mātam (probably 1928), <http://www.tamilcc.org/thamizham/ebooks/3/273/273.pdf>
11. **Kanakkadigaram 1958:** Koṟukkaiyūr Kāri Nāyaṇār, Tirunelvēli Tennintiya Caiva Cittānta Nūṟpatippuk Kaḷakam, Chennai, 1958
12. **Kanita Nuul:** Kaṇita Nūl, Institute of Asian Studies, Chennai, 1999
13. **Pope 1859:** A Tamil Handbook, 2nd edn, G U Pope, American Mission Press, Madras, 1859, <http://archive.org/details/atamilhandbook00popegoog>
14. **Pope AES:** A Tamil Handbook, 5th reprint, G U Pope, Asian Educational Services, Delhi, 1993
15. **Shuddhananda Bharati:** Iniya Tamil Ilakkaṇam, Yoki Śrī Cuttānanta Pāratiyār, Kavita Publications, Chennai, 1998
16. **Subramanian:** Paṇṭait Tamil Eḷuttukkaḷ, T N Subramanian, Madras Law Journal Press, Chennai, 1938
17. **Tamil Lexicon:** Tamil Lexicon, University of Madras, 1924-1936
18. **Wickremasinghe:** Tamil Grammar Self-Taught, M de Silva Wickremasinghe, Malborough and Co, London, 2nd edn, 1906, <http://books.google.com/books?id=9GNAAAAAYAAJ>
19. **Winslow:** A Comprehensive Tamil and English Dictionary, Miron Winslow, Madras, 1862

§9. Technical information

§9.1. Entries for UnicodeData.txt (also attached to PDF)

0BDE;TAMIL TRADITIONAL CREDIT SIGN;So;0;ON;;;;;N;;;;;
0BDF;TAMIL TRADITIONAL NUMBER SIGN;So;0;ON;;;;;N;;;;;
0BFB;TAMIL RINGGIT SIGN;Sc;0;ET;;;;;N;;;;;
0BFC;TAMIL CURRENT SIGN;So;0;ON;;;;;N;;;;;
0BFD;TAMIL AND ODD SIGN;So;0;ON;;;;;N;;;;;
0BFE;TAMIL SPENT SIGN;So;0;ON;;;;;N;;;;;
0BFF;TAMIL TOTAL SIGN;So;0;ON;;;;;N;;;;;

11FC0;TAMIL FRACTION ONE THREE-HUNDRED-AND-TWENTIETH;No;0;L;;;1/320;N;;;;;
11FC1;TAMIL FRACTION ONE ONE-HUNDRED-AND-SIXTIETH;No;0;L;;;1/160;N;;;;;
11FC2;TAMIL FRACTION ONE EIGHTIETH;No;0;L;;;1/80;N;;;;;
11FC3;TAMIL FRACTION ONE SIXTY-FOURTH;No;0;L;;;1/64;N;;;;;
11FC4;TAMIL FRACTION ONE FORTIETH;No;0;L;;;1/40;N;;;;;
11FC5;TAMIL FRACTION ONE THIRTY-SECOND;No;0;L;;;1/32;N;;;;;
11FC6;TAMIL FRACTION THREE EIGHTIETHS;No;0;L;;;3/80;N;;;;;
11FC7;TAMIL FRACTION THREE SIXTY-FOURTHS;No;0;L;;;3/64;N;;;;;
11FC8;TAMIL FRACTION ONE TWENTIETH;No;0;L;;;1/20;N;;;;;
11FC9;TAMIL FRACTION ONE SIXTEENTH;No;0;L;;;1/16;N;;;;;
11FCA;TAMIL FRACTION ONE TENTH;No;0;L;;;1/10;N;;;;;
11FCB;TAMIL FRACTION ONE EIGHTH;No;0;L;;;1/8;N;;;;;
11FCC;TAMIL FRACTION THREE TWENTIETHS;No;0;L;;;3/20;N;;;;;
11FCD;TAMIL FRACTION THREE SIXTEENTHS;No;0;L;;;3/16;N;;;;;
11FCE;TAMIL FRACTION ONE FIFTH;No;0;L;;;1/5;N;;;;;
11FCF;TAMIL FRACTION ONE QUARTER;No;0;L;;;1/4;N;;;;;
11FD0;TAMIL FRACTION ONE HALF;No;0;L;;;1/2;N;;;;;
11FD1;TAMIL FRACTION THREE QUARTERS;No;0;L;;;3/4;N;;;;;
11FD2;TAMIL FRACTION DOWNSCALING FACTOR KIIZH;No;0;L;;;1/320;N;;;;;
11FD3;TAMIL SIGN NEL;So;0;ON;;;;;N;;;;;
11FD4;TAMIL SIGN SUVADU;So;0;ON;;;;;N;;;;;
11FD5;TAMIL SIGN AAZHAAKKU;So;0;ON;;;;;N;;;;;
11FD6;TAMIL SIGN UZHAKKU;So;0;ON;;;;;N;;;;;
11FD7;TAMIL SIGN URI;So;0;ON;;;;;N;;;;;
11FD8;TAMIL SIGN MUUVUZHAKKU;So;0;ON;;;;;N;;;;;
11FD9;<reserved>
11FDA;TAMIL SIGN KURUNI;So;0;ON;;;;;N;;;;;
11FDB;TAMIL SIGN PDAKKU;So;0;ON;;;;;N;;;;;
11FDC;TAMIL SIGN MUKKURUNI;So;0;ON;;;;;N;;;;;
11FDD;TAMIL SIGN TUUNI;So;0;ON;;;;;N;;;;;
11FDE;TAMIL SIGN KALAM;So;0;ON;;;;;N;;;;;
11FDF;TAMIL SIGN PAISAA;Sc;0;ET;;;;;N;;;;;
11FE0;TAMIL SIGN ANAA;Sc;0;ET;;;;;N;;;;;
11FE1;TAMIL SIGN KAASU;Sc;0;ET;;;;;N;;;;;
11FE2;TAMIL SIGN PANAM;Sc;0;ET;;;;;N;;;;;
11FE3;TAMIL SIGN PON;Sc;0;ET;;;;;N;;;;;
11FE4;TAMIL SIGN VARAAGAN;Sc;0;ET;;;;;N;;;;;
11FE5;TAMIL SIGN BAARAM;So;0;ON;;;;;N;;;;;
11FE6;TAMIL SIGN GEJAM;So;0;ON;;;;;N;;;;;
11FE7;TAMIL SIGN KUZHI;So;0;ON;;;;;N;;;;;
11FE8;TAMIL SIGN VELI;So;0;ON;;;;;N;;;;;
11FE9;TAMIL WET CULTIVATION SIGN;So;0;ON;;;;;N;;;;;
11FEA;TAMIL DRY CULTIVATION SIGN;So;0;ON;;;;;N;;;;;
11FEB;TAMIL LAND SIGN;So;0;ON;;;;;N;;;;;
11FEC;TAMIL SALT PAN SIGN;So;0;ON;;;;;N;;;;;
11FED;TAMIL CHARGE SIGN;So;0;ON;;;;;N;;;;;
11FEE;TAMIL FIRST SIGN;So;0;ON;;;;;N;;;;;
11FEF;TAMIL SIGN MUDALIYA;So;0;ON;;;;;N;;;;;
11FF0;TAMIL SIGN VAGAIYARAA;So;0;ON;;;;;N;;;;;
11FF1;TAMIL SIGN CIRANJIIVI;So;0;ON;;;;;N;;;;;
11FF2;TAMIL SIGN PILLAI;So;0;ON;;;;;N;;;;;
11FF3;TAMIL SIGN RAAJA;So;0;ON;;;;;N;;;;;
11FF4;TAMIL ABBREVIATION KKU;Lo;0;L;;;N;;;;;
11FF5;TAMIL ABBREVIATION YUM;Lo;0;L;;;N;;;;;
11FF6;TAMIL ABBREVIATION VUM;Lo;0;L;;;N;;;;;
11FF7;TAMIL PUNCTUATION END OF TEXT;Po;0;L;;;N;;;;;

§9.2. Entries for NamesList.txt (also attached to PDF)

i) After:

0BD7 TAMIL AU LENGTH MARK

include:

@ Tamil symbols
0BDE TAMIL TRADITIONAL CREDIT SIGN
= varavu
* this is the traditional credit sign
x (tamil credit sign - 0BF7)
0BDF TAMIL TRADITIONAL NUMBER SIGN
= enn
* this is the traditional number sign
x (tamil number sign - 0BFA)

ii) After:

@ Tamil numerics

include:

@+ Tamil fractions are encoded at 11FC0-11FD2

iii) After:

@ Tamil symbols

include:

@+ More symbols are encoded in the Tamil Supplement block 11FC0-11FFF starting 11FD3

iv) Modify:

0BF3 TAMIL DAY SIGN
= naal

to:

0BF3 TAMIL DAY SIGN
= naal
= naazhi / padi
* denotes a measure of grain that equals 2 uri or 4 uzhakku
x (tamil sign uzhakku - 11FD6)
x (tamil sign uri - 11FD7)
= pillaiyaar suzhi
* denotes auspiciousness

v) Modify:

0BF7 TAMIL CREDIT SIGN
= varavu

to:

0BF7 TAMIL CREDIT SIGN
= eduppu
* denotes incoming cash which is set aside for unknown expenses
* sometimes used as the credit sign
* the traditional credit sign is different
x (tamil traditional credit sign - 0BDE)

vi) Modify:

0BFA TAMIL NUMBER SIGN
= enn

to:

0BFA TAMIL NUMBER SIGN
= niluvai
* denotes balance
* sometimes used as the number sign
* the traditional number sign is different
x (tamil traditional number sign - 0BDF)

and after that include:

@ Currency symbol
0BFB TAMIL RINGGIT SIGN
@ Tamil symbols
0BFC TAMIL CURRENT SIGN
= naaladu
0BFD TAMIL AND ODD SIGN
= silvaanam / sillarai
* not to be confused with the sign for "ciranjiivi"
x (tamil sign ciranjiivi - 11FF0)
0BFE TAMIL SPENT SIGN
= poga
* not to be confused with the abbreviation for "pillai"
x (tamil sign pillai - 11FF1)
0BFF TAMIL TOTAL SIGN
= aaga

vii) For the SMP block:

@@ 11FC0 Tamil Supplement 11FFF
@ Fractions
11FC0 TAMIL FRACTION ONE THREE-HUNDRED-AND-TWENTIETH
= mundiri
11FC1 TAMIL FRACTION ONE ONE-HUNDRED-AND-SIXTIETH
= araikkaani
11FC2 TAMIL FRACTION ONE EIGHTIETH
= kaani
11FC3 TAMIL FRACTION ONE SIXTY-FOURTH
= kaalviisam
11FC4 TAMIL FRACTION ONE FORTIETH
= araimaa
11FC5 TAMIL FRACTION ONE THIRTY-SECOND
= araiviisam
11FC6 TAMIL FRACTION THREE EIGHTIETHS
= mukkaani
11FC7 TAMIL FRACTION THREE SIXTY-FOURTHS
= mukkaalviisam
11FC8 TAMIL FRACTION ONE TWENTIETH
= maa
11FC9 TAMIL FRACTION ONE SIXTEENTH
= viisam / maakaani
11FCA TAMIL FRACTION ONE TENTH
= irumaa
11FCB TAMIL FRACTION ONE EIGHTH
= araikkaal
11FCC TAMIL FRACTION THREE TWENTIETHS
= mummaa
11FCD TAMIL FRACTION THREE SIXTEENTHS
= muuviisam / mummaamukkaani
11FCE TAMIL FRACTION ONE FIFTH
= naalumaa
11FCF TAMIL FRACTION ONE QUARTER
= kaal
11FD0 TAMIL FRACTION ONE HALF
= arai
11FD1 TAMIL FRACTION THREE QUARTERS
= mukkaal

11FD2 TAMIL FRACTION DOWNSCALING FACTOR KIIZH
 * when prefixed to a fraction, reduces its value by a factor of 1/320

@ Measures of grain

11FD3 TAMIL SIGN NEL
 * one grain of paddy

11FD4 TAMIL SIGN SUVADU
 * equals 360 nel

11FD5 TAMIL SIGN AAZHAAKKU
 * equals 5 suvadu

11FD6 TAMIL SIGN UZHAKKU
 * equals 2 azhakku

11FD7 TAMIL SIGN URI
 * equals 2 uzhakku

11FD8 TAMIL SIGN MUUVUZHAKKU
 * equals 3 uzhakku

11FD9 <reserved>
 * for the measure naazhi/padi that equals 2 uri or 4 uzhakku, use 0BF3
 x (tamil day sign - 0BF3)

11FDA TAMIL SIGN KURUNI
 * equals 8 naazhi/padi
 * also known as marakkal

11FDB TAMIL SIGN PADAKKU
 * equals 2 kuruni

11FDC TAMIL SIGN MUKKURUNI
 * equals 3 kuruni

11FDD TAMIL SIGN TUUNI
 * equals 2 padakku or 4 kuruni

11FDE TAMIL SIGN KALAM
 * equals 3 tuni
 * always ligates with a preceding 0BE7 tamil digit one

@ Old currency symbols

11FDF TAMIL SIGN PAISAA
 * old paisa comprises 3 pai and equals 1/64 of a rupee
 * new or naya paisa equals 1/100 of a rupee

11FE0 TAMIL SIGN ANAA
 * equals 4 old paisa

11FE1 TAMIL SIGN KAASU
 * alternate old currency

11FE2 TAMIL SIGN PANAM
 * equals 80 kaasu

11FE3 TAMIL SIGN PON
 * equals 10 panam

11FE4 TAMIL SIGN VARAAGAN
 * equals 42 panam
 * equals 3½ rupee, 3¾ as per some

@ Symbols of weight, length and area

11FE5 TAMIL SIGN BAARAM
 * approximately equals 500 pounds

11FE6 TAMIL SIGN GEJAM
 * approximately equals 3 feet

11FE7 TAMIL SIGN KUZHI
 * equals 16 square gejam

11FE8 TAMIL SIGN VELI
 * equals 2000 kuzhi

@ Agricultural symbols

11FE9 TAMIL WET CULTIVATION SIGN
 = nansey

11FEA TAMIL DRY CULTIVATION SIGN
 = punsey

11FEB TAMIL LAND SIGN
 = nilam

11FEC TAMIL SALT PAN SIGN
 = alam

@ Other symbols and abbreviations

11FED TAMIL CHARGE SIGN
 = vasam
 * possession

11FEE TAMIL FIRST SIGN
 = mudal

11FEF TAMIL SIGN MUDALIYA
 = et cetera
 * indicates items in a series

11FF0 TAMIL SIGN VAGAIYARAA
 = et cetera

* indicates items of a family or kind
 11FF1 TAMIL SIGN CIRANJIIVI
 * a polite title for a young person
 * not to be confused with the sign for "and odd"
 x (tamil and odd sign - 0BFD)
 11FF2 TAMIL SIGN PILLAI
 * not to be confused with the sign for "spent"
 x (tamil spent sign - 0BFE)
 11FF3 TAMIL SIGN RAAJA
 * in the phrase mahaa raaja raaja shrii
 * 0BF1 one hundred is often used as a substitute
 @ Abbreviations of word-endings
 11FF4 TAMIL ABBREVIATION KKU
 * used especially with numerals for the dative suffix
 11FF5 TAMIL ABBREVIATION YUM
 * and
 11FF6 TAMIL ABBREVIATION VUM
 * and
 @ Tamil punctuation
 11FF7 TAMIL PUNCTUATION END OF TEXT
 x (sinhala punctuation kunddaliya - 0DF4)

§9.3. Remarks on the choice of character properties

§9.3.1. Block Name

The chosen block name is “Tamil Supplement” in line with the existing block names Latin Supplement, Cyrillic Supplement and so on.

§9.3.2. Codepoints

The six characters used in clerical purposes listed in §3.4.3 are chosen to be placed in the BMP Tamil block as it already contains other clerical use characters from 0BF6 DEBIT SIGN upto 0BFA NUMBER SIGN. Of these two (TRADITIONAL CREDIT SIGN and TRADITIONAL NUMBER SIGN) need to be distinguished from their already encoded counterparts at 0BF7 and 0BFA and are hence placed in a different column (at 0BDE and 0BDF).

The RINGGIT SIGN is also proposed in the BMP block as it is akin to 0BF9 RUPEE SIGN. The codepoint chosen is 0BFB to be as close as possible to the latter.

The rest of the characters are placed in the SMP block in the order of fractions, measures, currency (all in ascending order of value) and miscellanea.

§9.3.3. Names

Names are modeled on existing character names, especially those in the Indic scripts.

The fractions are after the pattern <SCRIPT> FRACTION <NUMERATOR> <DENOMINATOR>S. When the denominator is formed by multiple words (a first in Unicode) we have used hyphens to join up those words for readability. This is permitted by the naming rules.

To be noted is the idiosyncratic name of 11FD2 TAMIL FRACTION DOWNSCALING FACTOR KIIZH. DOWNSCALING FACTOR indicates the purpose. The native term “kiizh” (kīl) is chosen for the name for simplicity rather than writing the value $1/_{320}$ out in words.

For the symbols, when the meaning or significance of the character is translatable, the format is <SCRIPT> <MEANING> SIGN. as in TAMIL DEBIT SIGN etc. So we propose TAMIL BALANCE SIGN etc. When it is not translatable, the format is <SCRIPT> SIGN <NATIVE_NAME>. as in DEVANAGARI SIGN AVAGRAHA etc. So we propose TAMIL SIGN NEL etc. To be noted is that 11FEE MUDALIYA and 11FEF VAGAIYARAA both are approximately translatable as “et cetera” but their distinction is not translatable. Further, in transliterating the native names, we use the popular conventions (such as zh for ழ, etc) for readability by native users.

Three ligature-like abbreviations of word endings are explicitly labeled as abbreviations and for their sound value: TAMIL ABBREVIATION KKU/YUM/VUM. It is requested to not label these as letters as they are not recognized as letters of the Tamil alphabet.

§9.3.4. General Category and Bidi

The fractions are GC=No, the currency symbols GC=Sc, the abbreviated word endings GC=Lo, the punctuation mark GC=Po, and the rest GC=So. The appropriate Bidi values are allotted.

§9.3.5. Numerical value

The fractions are allotted appropriate numerical values. The special character 11FD2 TAMIL FRACTION DOWNSCALING FACTOR KIIZH is allotted the value 1/320.

§9.4. Collation, other properties and security

The fractions are encoded in ascending order of value. They should be collated that way. The 11FD2 TAMIL FRACTION DOWNSCALING FACTOR KIIZH கீ should be sorted before the least fraction 1/320 (even though it is equal in value to it), so that other fractions prefixed by it will be collated before the least fraction. That is:

கீ வு < கீ ற < ... < கீ வ < கீ ற < கீ ள < வு < ற < ... < வ < ற < ள < க < உ ...

The measures are also encoded in ascending order. To be noted is however that 11FD8 < 0BF3 < 11FDA since the measure nāḷi / paṭi is not disunified from 0BF3.

The three word-ending abbreviations should be sorted as if they were written out fully, but after the explicitly written forms. Thus tīyum < tī{yum} < tīyul and so on.

Linebreak and other properties may be allotted as of other similar Indic characters.

These are all historical use characters being encoded for the purpose of storing old documents accurately as digital text. As such, they do not need to be used in IDNs in the contemporary context and hence they may be prohibited from IDNs altogether to avoid any potential for security issues, just as the existing Tamil numerals are already prohibited from IDNs (Mani Manivannan, email communication dt 2010-May-12).

§9.5. Font issues

The glyphs for the proposed characters have been designed by the present author largely based on the glyphs of the existing Lohit Tamil font (<https://fedorahosted.org/lohit/>) under derivative rights provided by the Open Font Licence (<http://scripts.sil.org/OFL>, hereafter OFL) under which the Lohit fonts are licensed. Where it was not possible, glyphs were designed by the author from scratch based on the obtained attestations.

The author has donated all these glyphs (both derived and self-designed) back to the Lohit project under the conditions of the OFL. The Lohit fonts maintainers, especially Pravin Satpute of Red Hat, have graciously accepted this contribution. My thanks are due to them. (See https://bugzilla.redhat.com/show_bug.cgi?id=839303.)

It is requested to utilize these glyphs for the code chart for the additional characters proposed. In addition, it is requested that the entire Tamil code chart may be updated to use the glyphs from Lohit Tamil to ensure stylistic consistency among all the encoded Tamil characters in the Tamil and Tamil Supplement blocks. Lohit Tamil is a font of good quality and its maintainers have affirmed that it may be freely used for the Unicode code chart under the terms of the OFL (as can be seen in the bugzilla page linked above).

Pending encoding of these characters, a temporary version of the Lohit Tamil font version 2.5.1 and its Classical variant is uploaded to <http://pravins.fedorapeople.org/tamil-fraction-symbol-proposal-fonts/> with the additional glyphs mapped to the PUA. This font may be used by those desiring to use the additional characters' glyphs until such date as they are officially encoded. The same location also contains the Lohit Tamil Chart font to be eventually used for the code charts containing merely the glyphs for the existing and proposed characters mapped to their appropriate codepoints.

*

(go to next page)

§10. Code Charts

	0B8	0B9	0BA	0BB	0BC	0BD	0BE	0BF
0		ஐ		ர	ீ	ஓ		ய
1				ற	ு			ள
2		ஓ		ல	ூ			த
3	ஃ	ஓ	ண	ள				உ
4		ஓள	த	ழ				ம்
5	அ	க		வ				ஶ
6	ஆ			ய	ெ		ஃ	பு
7	இ			ஷ	ே	ள	க	ஶ
8	ஈ		ந	ஸ	ை		உ	ஷ
9	உ	ங	ன	ஹ			ந	ஶ
A	ஊ	ச	ப		ொ		ச	நீ
B					ோ		ரு	நீ
C		ஐ			ெள		கூ	நு
D					ஃ		எ	ஶ
E	எ	ஞ	ம	ா		வூ	அ	பீ
F	ஏ	ட	ய	ி		ஊ	கூ	ஆ

Tamil Supplement

11FC 11FD 11FE 11FF

0	வத	உ	அணு	வூ
1	ட	த	கூ	சூ
2	லு	சூ	பு	பூ
3	லூ	று	லூ	ரூ
4	சு	ஸூ	வூ	றூ
5	சூ	று	று	யூ
6	சூ	த	கூ	வூ
7	சூ	வரி	கூ	வூ
8	ப	த	லூ	
9	ய		தூ	
A	ஃ	ங	பஹு	
B	ஹு	ஹு	நூ	
C	ஹ	ஹ	ஹு	
D	ஹ	த	ஹு	
E	ஹ	ள	ஹு	
F	வ	ஹ	ஹு	

§11. Official Proposal Summary Form

(Based on N3902-F)

A. Administrative

1. Title

Proposal to encode Tamil fractions and symbols

2. Requester's name

Shriramana Sharma

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

Individual Contribution

4. Submission date

2012-Jul-17

5. Requester's reference (if applicable)

6. Choose one of the following: This is a complete proposal (or) More information will be provided later

This is a complete proposal.

B. Technical – General

1. Choose one of the following:

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

No

1b. The proposal is for addition of character(s) to an existing block, Name of the existing block

These characters are additions to the Tamil encoding, but they cannot all be encoded in the existing BMP Tamil block due to lack of space. As a result, a new Tamil Supplement block is requested.

2. Number of characters in proposal

62 (sixty two)

3. Proposed category

Category B1, specialized small

4. Is a repertoire including character names provided?

Yes

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

Yes

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

Yes

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?

Shriramana Sharma

b. Identify the party granting a license for use of the font by the editors (include address, e-mail etc.)

Shriramana Sharma and other contributors to the Lohit Tamil font, under derivative rights granted by the OFL. See <https://fedorahosted.org/lohit/>.

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.

See detailed proposal.

C. Technical – Justification

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

See §2 of this document for details. The present document seeks to be a finalized proposal for all attested Tamil fractions and symbols.

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?

G Balachandran of ICTA Sri Lanka, various Tamil scholars participating in the C-Tamil mailing list (ctamil-at-services.cnrs.fr), some members of INFITT WG02. See §1 of this document for details.

2c. If YES, available relevant documents

The matter was largely discussed in person or via email.

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

Those who desire to store as digital text old Tamil manuscripts involving these characters, and those who may desire to revive the use of at least some of these characters

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

Rare

4b. Reference

See detailed proposal.

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

Scholars who work with manuscripts will use these characters.

5b. If YES, where?

Largely in research institutions around the world involved with Tamil and some Grantha manuscripts.

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

No

6b. If YES, is a rationale provided?

6c. If YES, reference

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

Yes

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?

Some characters may be similar to existing Tamil/Grantha letters or ligatures thereof.

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

Yes

10c. If YES, reference

The proposal characters' properties would be different, i.e. GC=No, GC=So etc.

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

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?

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?

13c. If YES, reference:

-O-O-O-