



INFORMATION TECHNOLOGY
DEPARTMENT,
SECRETARIAT,
Chennai- 600 009.

Letter No. O-17/116/2016-IT (e.Govt), dated: 15.03.2017

From
Thiru T.K. Ramachandran, IAS.,
Principal Secretary to Government.

To
Ms. Lisa Moore
Chair, Unicode Technical Committee
Vice President, Chief Financial Officer,
The Unicode consortium
P.O.Box No. 391476 Mountain view,
CA 94039-1476, USA. (w.e)

Madam,

Sub: Tamil Fractions and special symbols in Unicode -
Revised – Finalized proposal sent - Regarding.

Ref: Tamil fractions and symbols encoding proposal –
Extended feedbacks – L2/16-062, Dated 25.02.2016.

I am to invite your attention to the reference cited and to state that the Tamil Virtual Academy (TVA) functioning under the Government of Tamil Nadu has forwarded an amendment proposal and the same was published in the Unicode document register by UTC on 25.02.2016 through L2/16-062 for feedback. The revised proposal L2/16-062 received a few comments from Sri Ramana Sharma, who has sent the proposal L2/12-231, and from a few others. After receiving the comments the expert committee constituted for this purpose including Sri Ramana Sharma met on the above subject and finalized the proposal. The revised and finalized proposal on Tamil fractions and symbols is enclosed for your consideration.

2. Since the time available is short, I am to request you to take up the matter immediately with UTC and ISO and consider the proposal favourably.

Yours faithfully,


15/3/2017

for Principal Secretary to Government.

15/3/2017

Copy to:
The Director,
TVA, Anna University Campus,
Gandhi Mandapam Road, Kottur, Chennai – 600 025.

Finalized proposal to encode Tamil fractions and symbols

Government of Tamil Nadu

2017-Jan-21

§1. Introduction



Based on the initial Tamil Fractions and Symbols proposal L2/12-231 by Shriramana Sharma, alternate proposal L2/16-062 by Government of Tamil Nadu and feedback L2/16-114 by Sharma, the Government of Tamil Nadu organized a committee including subject experts as well as Shriramana Sharma and a meeting was held on 2016-Oct-20 at Tamil Virtual Academy, Chennai at 11 AM. The meeting arrived at a consensus for fifty one (51) characters to be encoded in the Tamil Supplement block of the Supplementary Multilingual Plane (SMP) of Unicode from 11FC0 to 11FFF as in the existing roadmap. This is the formal proposal for these characters to be added to the existing Tamil Unicode encoding.

In this context it is noted that:

- 1) The last status of the previous proposal was reflected in the draft additional repertoire for ISO/IEC 10646:2016 5th edition CD2 L2/15-339 WG2 N4705 pp 94-96.
- 2) Based on the disposition of the comments received from the India National Body as reflected in L2/16-091 WG2 N4713 pp 16-17, the Tamil Supplement block has now been totally removed from the ISO ballot.
- 3) The Unicode pipeline page <http://www.unicode.org/alloc/Pipeline.html> still (as of 2017-Jan-21) reflects the ISO CD2. It is presumed that this needs to be updated.
- 4) For the sake of simplicity, it is recommended to start over with a clean slate and thus a fresh proposal is made for the entire set of characters without any reference to the ISO CD2 names, codepoints etc since those are no longer valid.




















Attestations of the characters are summarized in this document for convenience of review. (For some historic shapes, due to the difficulty of getting direct attestations, the shapes are based on subject experts' recommendations.) Detailed discussions are only to be found in the earlier proposals L2/12-231 by Shriramana Sharma ("Sharma") and L2/16-062 by the Government of Tamil Nadu ("GoTN") and are not duplicated here for brevity and clarity.

The recommendation to disunify some characters which were not disunified earlier is also briefly discussed in this document. The minutes of the Chennai meeting of 2016-Oct-20 are forwarded separately for information of the committees.

Finally, in this document the glyph of *ceviṭu* has been corrected from  seen in the above documents to the actual shape  seen in the attestations.

§2. List of characters

§2.1. Fractions

	Glyph	x/320	Value	Words	Tamil word
1.		1/320	1/320	one three-hundred-and-twentieth	muntiri
2.		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.		8/320	1/40	one fortieth	araimā
6.		10/320	1/32	one thirty-second	araivīcam
7.		12/320	3/80	three eightieths	mukkāṇi
8.		15/320	3/64	three sixty-fourths	mukkālvīcam
9.		16/320	1/20	one twentieth	mā
10.		20/320	1/16	one sixteenth	vīcam / mākāṇi
11.				<alternate form of above>	
12.		32/320	1/10	one tenth	irumā
13.		40/320	1/8	one eighth	araikkāl
14.		48/320	3/20	three twentieths	mummā
15.		60/320	3/16	three sixteenths	mūvīcam / mummāmukkāṇi
16.		64/320	1/5	one fifth	nāṇkumā
17.		80/320	1/4	one quarter	kāl
18.		160/320	1/2	one half	arai
19.				<alternate form of above>	

20. ௩௪ 240/320 3/4 three quarters mukkāl

Fractions prefix:

21. ௧௩ 1/320 × 1/320 multiply by 1/320 kīl

§2.2. Measures of grain

Glyph	Meaning	Tamil word
22. ௪௩	1 grain of paddy	nel
23. ௪௪	360 nel	ceviṭu
24. ௪௫	5 cuvaṭu	ālākku
25. ௪௬	2 ālākku	uḷakku
* ௪௭ (SEQUENCE)	2 uḷakku	uri
26. ௪௮	3 uḷakku	mūvuḷakku
* ௪௯ (EXISTS AT 0BF3)	2 uri (4 uḷakku)	nāli / paṭi
27. ௪௮	8 nāli / paṭi	kurun̄i / marakkāl
* ௪௮ (EXISTS AT 0B99)	<alternate form of above>	
28. ௪௮	2 kurun̄i	patakku
29. ௪௮	3 kurun̄i	mukkurun̄i
* ௪௮ (EXISTS AT 0BA4)	2 patakku (4 kurun̄i)	tūṇi
* ௪௮ (EXISTS AT 0BB3)	3 tūṇi	kalam

§2.3. Currency

Glyph	Meaning	Tamil word
30. ௪௮	small currency unit	kācu
31. ௪௮	“money”	paṇam
32. ௪௮	gold coin	pon
33. ௪௮	gold coin bearing boar insignia	varākaṇ

§2.4. Weight and area

Glyph	Meaning	Tamil word
34. ௪௮	≅ 500 lb. (227 kg.)	pāram

35.	கூல்	1 sq. kōl where 1 kōl \cong 11 ft.	kuḷi
36.	வீல்	2000 kuḷi	vēli

§2.5. Agriculture

37.	நீர்	wet cultivation	nañcey
38.	பூர்	dry cultivation	pun̄cey
39.	நீர்	land	nilam
40.	அப்பா	salt pan	uppaḷam

§2.6. Clerical

41.	வரவு	credit	varavu
42.	நா	number	eṇ
43.	நிலை	current	nāḷatu
44.	சீர்	... and odd	cilvāṇam / cillarai
45.	பீர்	... having been spent	pōka
46.	அகா	total	āka

§2.7. Other

47.	வா	in possession	vacam
48.	மு	starting from	mutal
49.	மு	et cetera (in a series)	mutaliya
50.	வா	et cetera (of a kind)	vakaiyaṛā

§2.8. Punctuation

51.	ஂ	end of text	
-----	---	-------------	--

The above glyphs of the characters have been checked by subject experts based on the original sources as provided in the earlier proposals L2/12-231 and L2/16-062.

§3. Disunification

1. For the fraction one twentieth, the initial proposal had proposed to encode a separate character ூ glyphically identical to the existing 0BAA letter PA. Based on the glyphic identity, the UTC had recommended to not disunify as noted in L2/13-047 p 1. However the subject experts identify a rounded ு as the original glyph in contrast with the angular 0BAA ூ and thus recommend to disunify.
2. Likewise for the fraction one fourth, the experts identify ௃ as the original glyph in contrast with the lower-headed 0BB5 ெ.
3. Likewise for the measure *kurun̄i/marakkāl*, the experts identify ே as the original glyph but acknowledge that the distinct 0B99 ை is seen in later usage.
4. Likewise for the measure *ulakku*, the experts identify ொ as the original glyph in contrast with the looped form ோ proposed for the fraction three quarters.
5. The initial proposal had not proposed to disunify the two forms of the fraction one sixteenth ் (then shown as ்) and ். The experts feel that the two are historically used distinctly though they have the same connotation and it would be useful to be able to distinguish them in plain text.
6. Likewise for the two forms of the fraction one half: ் and ்.

§4. Attestations

The character numbers below refer to the list in §2 above.

#1,2,3: Sharma, p 29

வகு	முத்திரை	is the three hundred and twentieth	
	part of the Integer,		$\frac{1}{320}$
ந	அரைக்காணி		$\frac{1}{160}$ $\frac{2}{320}$
சி	காணி		$\frac{1}{80}$ $\frac{4}{320}$

#4: Sharma, p 38

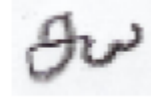
ஐ கால் வீசம் (ஐஊ) $\frac{1}{64}$

#5: Sharma, p 29

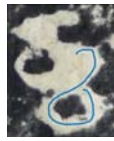
சு அரோமா - - - - - 1/40 8/320

#6: GoTN, p 24

the right symbol



#7: GoTN, p 13



#8: GoTN, p 25

the shape (SII-Vol.10, Inscription# 174)

#9: GoTN, p 16



#10: GoTN, p 9

is to be changed to



#11: GoTN, pp 10,11



#12: Sharma, p 29

இ இரண்டுமா or இருமா - - $\frac{1}{10} - \frac{33}{320}$

#13: Sharma, p 36

அரைக்கால் ஹு

#14: Sharma, p 29

ந மும்மா or முன்றுமா - - $\frac{3}{20} - \frac{48}{320}$

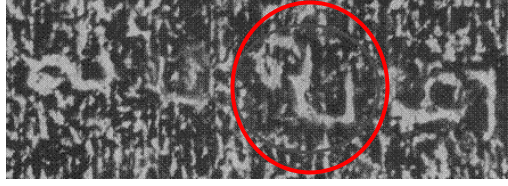
#15: GoTN, p 8

தரு: அகம்மை
புக்க ஒன்பதி
[ஓ]ள[ஸம்]ம
ராகம் [ஓ]ள[ஸம்]
[க]நம் உயி
[க]கால் [ச]

#16: GoTN, p 22

4 சுக் இன்னிலம் ந
புத்தடி நு சுரு-ல் நாலுமாவரையும்

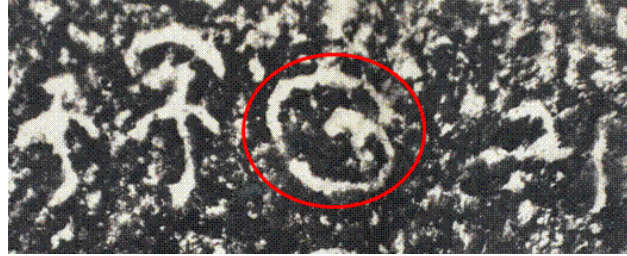
#17: GoTN, p 15



#18: Sharma, p 29

உ அரை - - - - - $\frac{1}{2}$ — $\frac{160}{320}$

#19: GoTN, p 19



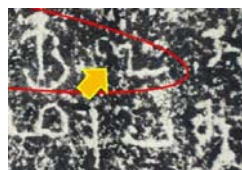
#20: Sharma, p 29

ந முக்கால் - - - - - $\frac{5}{4}$ — $\frac{240}{320}$

#21: Sharma, p 28

இதுவுமது.
கூ-ல்-க க்ருப்பெபான்-உய-க்ரும்-நகரிவநகீ-கி க்ருமாற,உய
ந-யடு-உய-கி-ச-உய-ரி-ஹ-உய-வந-படு-உய-கீ-கிடு-ஆ யககி
ஆதலால் ஆயிலென்றுடையான்-யககி-பொன்னிறப்பாணென்று
சொல்லது.

#22: GoTN, p 26



#23,24: Sharma, p 39



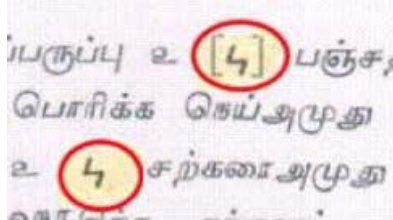
#25: GoTN, p 31

சூ for UZHAKKU

#26: GoTN, p 33

the form of ருது

#27: GoTN, p 31



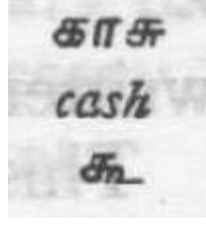
#28: GoTN, p 32

ருது in the primary resources

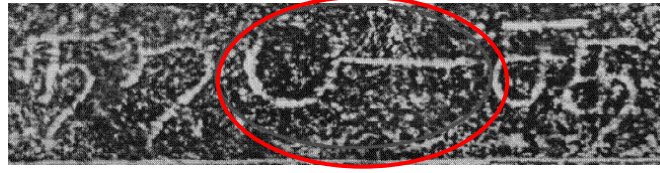
#29: Sharma, p 40



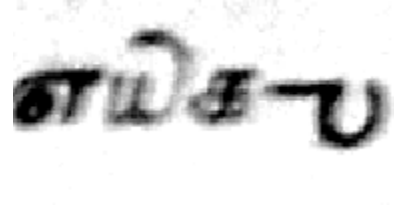
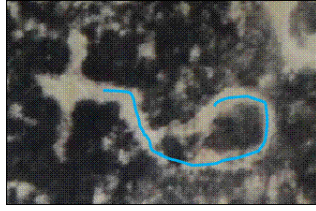
#30: Sharma, p 49



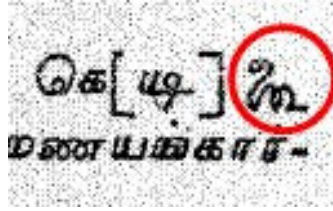
#31: GoTN, p 34



#32: GoTN, p 37



#33: GoTN, p 39



#34: Sharma, p 45

பாரம்..... ஸூ

#35: Sharma, p 46

கூழி..... கூ.

#36: Sharma, p 52

வேலி லி

#37: Sharma, p 49

நன்செய்
wet cultivation
நு

#38: Sharma, p 48

புன்செய்
dry cultivation
பு

#39: GoTN, p 45



#40: Sharma, p 50

ஆ = அளம்

#41: Sharma, p 51

வூ = வரவு

#42: Sharma, p 50

நு = தம்பர்

#43: Sharma, p 54

பித்விநுமீ (கூவ)
போது சாகுபடிக

#44: Sharma, p 50

சு = சில்வானம்

#45: Sharma, p 45

போக..... டு

#46: Sharma, p 54

அடி மொத்தம் உதாரண சாத்திரங்களில்
தனிப்பெயராய் வந்த சாத்திரங்கள்

#47: Sharma, p 45

வசம்..... ஸு

#48: GoTN, p 47

காவது முடி வறெழையில்

#49: Sharma, p 53

நய	-	-	முப்பது	-	-	-	-	-	-	30
நயத	-	-	முப்பத்தொன்று	மி	-	-	-	-	-	31 &c.
சய	-	-	நாற்பது	-	-	-	-	-	-	40
சயஉ	-	-	நாற்பத்திரண்டு	மி	-	-	-	-	-	42 &c.

#50: Sharma, p 48

வகையரா
belonging to
உல

#51: Sharma, p 54

இததாலுரியபபனுனுகிரதாவது ஸ

§5. Unicode Character Properties

The additions to UnicodeData.txt for the proposed SMP Tamil Supplement block are inlined below. These and NamesList.txt entries for the same are also provided as attachments to this PDF. It is noted that the current NamesList.txt for the existing BMP Tamil block contains some references which are invalidated by this proposal. As such, an updated NamesList.txt section for the same including some spelling corrections is also attached.

```
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-1;No;0;L;;;;1/16;N;;;;;
11FCA;TAMIL FRACTION ONE SIXTEENTH-2;No;0;L;;;;1/16;N;;;;;
11FCB;TAMIL FRACTION ONE TENTH;No;0;L;;;;1/10;N;;;;;
11FCC;TAMIL FRACTION ONE EIGHTH;No;0;L;;;;1/8;N;;;;;
11FCD;TAMIL FRACTION THREE TWENTIETHS;No;0;L;;;;3/20;N;;;;;
11FCE;TAMIL FRACTION THREE SIXTEENTHS;No;0;L;;;;3/16;N;;;;;
11FCF;TAMIL FRACTION ONE FIFTH;No;0;L;;;;1/5;N;;;;;
11FD0;TAMIL FRACTION ONE QUARTER;No;0;L;;;;1/4;N;;;;;
11FD1;TAMIL FRACTION ONE HALF-1;No;0;L;;;;1/2;N;;;;;
11FD2;TAMIL FRACTION ONE HALF-2;No;0;L;;;;1/2;N;;;;;
11FD3;TAMIL FRACTION THREE QUARTERS;No;0;L;;;;3/4;N;;;;;
11FD4;TAMIL FRACTION DOWNSCALING FACTOR KIIZH;No;0;L;;;;1/320;N;;;;;
11FD5;TAMIL SIGN NEL;So;0;ON;;;;;N;;;;;
11FD6;TAMIL SIGN CEVITU;So;0;ON;;;;;N;;;;;
11FD7;TAMIL SIGN AAZHAAKKU;So;0;ON;;;;;N;;;;;
11FD8;TAMIL SIGN UZHAKKU;So;0;ON;;;;;N;;;;;
11FD9;TAMIL SIGN MUUVUZHAKKU;So;0;ON;;;;;N;;;;;
11FDA;TAMIL SIGN KURUNI;So;0;ON;;;;;N;;;;;
11FDB;TAMIL SIGN PATHAKKU;So;0;ON;;;;;N;;;;;
11FDC;TAMIL SIGN MUKKURUNI;So;0;ON;;;;;N;;;;;
11FDD;TAMIL SIGN KAACU;Sc;0;ET;;;;;N;;;;;
11FDE;TAMIL SIGN PANAM;Sc;0;ET;;;;;N;;;;;
11FDF;TAMIL SIGN PON;Sc;0;ET;;;;;N;;;;;
11FE0;TAMIL SIGN VARAAKAN;Sc;0;ET;;;;;N;;;;;
11FE1;TAMIL SIGN PAARAM;So;0;ON;;;;;N;;;;;
11FE2;TAMIL SIGN KUZHI;So;0;ON;;;;;N;;;;;
11FE3;TAMIL SIGN VELI;So;0;ON;;;;;N;;;;;
11FE4;TAMIL WET CULTIVATION SIGN;So;0;ON;;;;;N;;;;;
11FE5;TAMIL DRY CULTIVATION SIGN;So;0;ON;;;;;N;;;;;
11FE6;TAMIL LAND SIGN;So;0;ON;;;;;N;;;;;
11FE7;TAMIL SALT PAN SIGN;So;0;ON;;;;;N;;;;;
11FE8;TAMIL TRADITIONAL CREDIT SIGN;So;0;ON;;;;;N;;;;;
11FE9;TAMIL TRADITIONAL NUMBER SIGN;So;0;ON;;;;;N;;;;;
11FEA;TAMIL CURRENT SIGN;So;0;ON;;;;;N;;;;;
11FEB;TAMIL AND ODD SIGN;So;0;ON;;;;;N;;;;;
11FEC;TAMIL SPENT SIGN;So;0;ON;;;;;N;;;;;
11FED;TAMIL TOTAL SIGN;So;0;ON;;;;;N;;;;;
11FEE;TAMIL IN POSSESSION SIGN;So;0;ON;;;;;N;;;;;
11FEF;TAMIL STARTING FROM SIGN;So;0;ON;;;;;N;;;;;
11FF0;TAMIL SIGN MUTHALIYA;So;0;ON;;;;;N;;;;;
11FF1;TAMIL SIGN VAKAIYARAA;So;0;ON;;;;;N;;;;;
11FFF;TAMIL PUNCTUATION END OF TEXT;Po;0;L;;;;;N;;;;;
```

§6. Code chart

	11FC	11FD	11FE	11FF
0	ஹ	உ	ஹை	மு
1	ஈ	உ	ஹு	வூ
2	ஐ	ஓ	கூ	
3	ஹை	ஹை	ஹை	
4	சு	கூ	ஹு	
5	கூ	ஹை	பூ	
6	ஹை	ஹை	ஹை	
7	கூ	ஹை	ஹை	
8	ப	ஹை	ஹை	
9	ஹை	ஹை	ஹை	
A	ஹை	ஹை	ஹை	
B	ஹை	ஹை	ஹை	
C	ஹை	ஹை	ஹை	
D	ஹை	ஹை	ஹை	
E	ஹை	ஹை	ஹை	
F	கூ	ஹை	ஹை	ஹை

§7. Official Proposal Summary Form

(Based on N3902-F)

A. Administrative

1. Title

Finalized proposal to encode Tamil fractions and symbols

2. Requester's name

Government of Tamil Nadu

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

Member body

4. Submission date

2017-Jan-21

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

51 (fifty one)

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?

See previous proposals L2/12-231 and L2/16-062.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

See previous proposals L2/12-231 and L2/16-062.

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)?

See previous proposals L2/12-231 and L2/16-062.

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 previous proposals L2/12-231 and L2/16-062.

C. Technical – Justification

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

This document replaces L2/12-231, L2/13-047 and L2/16-062 as per the decision taken at the subject experts meeting organized by the Government of Tamil Nadu on 2016-Oct-20 at Chennai.

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?

Subject experts at the aforementioned meeting. Also see previous proposals L2/12-231 and L2/16-062.

2c. If YES, available relevant documents

Please refer to the minutes of the aforementioned meeting (to be submitted separately).

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 old Tamil epigraphs and manuscripts using these characters as digital text

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?

Yes

5b. If YES, where?

In limited research contexts of the old Tamil writings mentioned above.

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 are similar (but not identical) to existing Tamil letters.

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

Yes

10c. If YES, reference

They have a consistently distinct shape.

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-