Comment on L2/13-047: Revised Proposal to encode Tamil fractions and Symbols

Dr. Naga Ganesan (naa.ganesan@gmail.com)

1.0 This document has two sections: (I) To use consistent transliteration of character names and annotations for Tamil fractions and symbols, following Madras University Lexicon system which is used very widely in the last few decades, and (II) To request UTC to avoid having dual encoding (2 new code points) for ந்- Tamil Number Sign (U+0BFA) and எவ - Tamil Credit Sign (U+0BF7). These are glyph variants serving the same purpose, and are used interchangeably in Accounting and Business practice as attested in documents and confirmed by Accountants, Traditional *KaNakkappiLLai*-s, clerks, and scholars who study Tamil Manuscripts. Please note that Tamil credit sign and Tamil number sign has many alternate symbols, and they can be used in different fonts instead of giving atomic code points for every glyph variant, as they are used interchangeably. The comparison is with the Tamil OM sign which also has similarly alternate glyph variants which are used interchangeably, and note that Tamil OM sign glyph variants are not given multiple atomic code points. In the same way, three glyphs each for Tamil credit sign and Tamil number sign works instead three different atomic code points in Tamil script encoding in the Unicode standard.

Section I: Transliteration of Tamil words for Characters

During the last UTC meeting I requested standard Tamil transcribing for words like "*aaka*" and "*poka*", instead of "*aaga*" and "*poga*", annotations for U+0BFE and U+0BFF respectively. And, this has been done in the revised proposal, L2/13-047. Along similar lines, the modifications for the following words are requested.

Table 1. Transliterating Tamil words for fractions and symbols in L2/13-047

| Proposed Code Point In L2/13-047	Tamil word in L2/13-047	Requested Change to match with Tamil transliteration standard	Remark
11FD2	SUVADU	CUVADU	Character name
11FDD	VARAAGAN	VARAAKAN	Character name
11FDE	BAARAM	PAARAM	Character name
11FDF	GEJAM	KESAM	Character name
11FE9	VAGAIYARAA	VAKAIYARAA	Character name
0BF3	suzhi	cuzhi	Word in Annotation
0BF9	rupai	ruupaay	Word in Annotation
0BFD	silvaanam/sillarai	cilvaanam/cillarai	Word in Annotation
0BF6	patru	parru	Word in Annotation
0BF3	naal	naall	Word in Annotation

Note: For 0BF3, naal means "4", while naall means "day". Compare 'enn' annotation in 0BFA. Let us be consistent in character names and annotations in using standard transliteration of Tamil words in Latin script.

Section II: Glyph Variations of Tamil Credit Sign (U+0BF7) and Tamil Number Sign (U+0BFA):

There are basically just very few symbols in use in modern Tamil, and all of them are currently encoded already in the Unicode standard in Tamil code chart. INFITT has requested *a separate block* for the archaic and not-in-use symbols and fractions in SMP (Ref. L2/09-398, L2/10-408). As S. Sharma says in L2/12-231, pg. 7, "*most Tamilians today are totally unaware of these old written forms*." Hence, it makes sense to encode all these unused and archaic symbols in the SMP block.

It is well known that Tamil credit sign, வ (U+0BF7) is just one variant of the available glyphs which are used interchangeably in Tamil accounting practice for credit sign. The original proposal to encode U+0BF7, written almost a decade+ ago, was not a well-researched one as it does not show the glyph variations, nor does it provide any printed book references for the glyph variations. The printed references for the glyph of Tamil credit sign involves the Tamil letter, VA (வ) of the word, VARAVU just as Tamil debit sign, U+0BF6 involves the Tamil letter PA (ப) of the word, PARRU. Some of the printed books' evidence of Tamil credit sign are documented in L2/12-150: http://www.unicode.org/L2/L2012/12150-tamil-credit.pdf

Tamil debit and credit signs, showing letters PA and VA respectively (Dr. G. U. Pope, 1855)
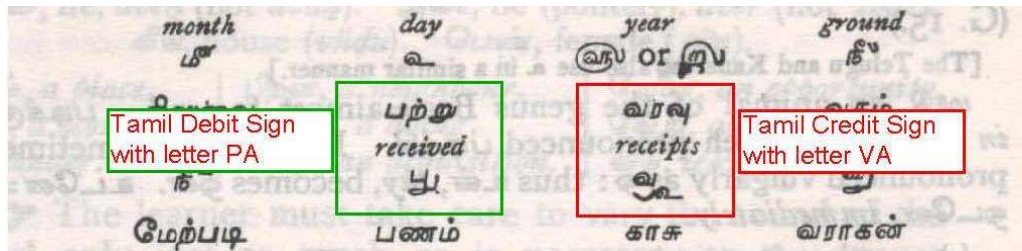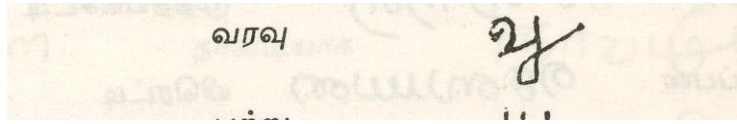

Dr. Wickramasinghe, London, 1906 (Note the difference in glyph shape)

Dr. R. Gruenendahl (Germany, 2001):

Two Tamil credit signs (used as alternate glyph variants interchangeably)
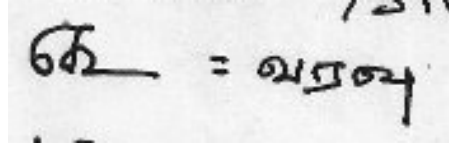
AavaNam, Journal of the Tamil Nadu Archaeological Society, Volume 8, 1997, Tamil University, Thanjavur, pg. 116

Note the letter VA (வ) as in first 3 glyphs. During the research on glyph variants of Tamil credit and number signs, the experts in Tamil Nadu recommend this glyph as the representative Tamil Credit sign for use in Unicode Tamil code chart, and the current glyph be replaced with this glyph. They point out the reason for their choice: Tamil credit sign glyph starting with letter VA (வ) letter just like Tamil debit sign starting with PA (ப) letter as we have in Tamil code chart of the Unicode standard.

Another glyph for Tamil credit sign. Reference: AavaNam, Journal of the Tamil Nadu Archaeological Society, Volume 8, 1997, Tamil University, Thanjavur, pg. 116

Research has been conducted in the field in Tamil Nadu with Tamil manuscriptologists, academics, Business clerks, KaNakkappillais, Accountants who advise that the Tamil credit sign and Tamil number sign (whose alternate glyphs are to be discussed next) are used interchangeably, and there is no difference in Accounting practice when they replace one glyph variant with the other glyph. The shapes of **SIX** equivalent glyph variants of Tamil credit sign are provided above. These 6 glyphs of the Tamil Credit Sign are like six glyphs of the Tamil OM sign. Obviously, atomic code points for each of the six glyph shapes of Tamil credit sign are not needed in Tamil code chart. Hence it is requested that atomic code point for each of
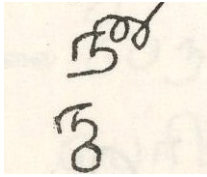
the equivalent 6 glyph variants are not encoded in the Unicode standard, and the glyphs are equivalent and alternately used. So, it is requested an additional atomic code point for Tamil credit sign in U+0BFC should not be encoded, and the glyph variation in Tamil credit sign can be handled easily in fonts. Comparison is with Tamil Om sign which has a wide glyph range, and the Unicode standard has just one atomic code point for it.

Similarly, let us look at the alternate glyph variations of Tamil number sign, U+0BFA which are used interchangeably. Tamil Number Sign (U+0BFA):

Reinhold Gruenendahl, 2001 book:

AavaNam, Journal of the Tamil Nadu Archaeological Society, Volume 8, 1997, Tamil University, Thanjavur, pg. 116:

     Two additional glyphs for Tamil Number sign (U+0BFA)

Hence it is requested that atomic code point for each of the equivalent 4 glyph variants of Tamil number sign are not encoded in the Unicode standard, and the glyphs are equivalent and alternately used. So, it is requested an additional atomic code point for Tamil credit sign in U+0BFB should not be encoded, and the glyph variations in Tamil number sign can be handled easily in fonts.

In sum,
(a) 6 glyph variants used alternately as equivalents for Tamil credit sign (U+0BF7) exist, and it is requested that dual encoding with atomic code point in U+0BFC for Tamil credit sign should not be done in Tamil encoding.

(b)) 4 glyph variants used alternately as equivalents for Tamil number sign (U+0BFA) exist, and it is requested that dual encoding with atomic code point in U+0BFB for Tamil number sign should not be done in Tamil encoding.

(c) All the symbols and fractions in L2/13-047 are archaic and not-in-use, and as per the original proposals, their place is together as one block in SMP. Arbitrarily encoding a few in BMP and others in SMP will create a bias among the archaic symbols. Hence it is requested all the fractions and symbols, when encoded, be moved into the SMP Tamil additional block.