

## Regarding my Tamil fractions/symbols proposal

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

2012-Jul-17

I request clarification from the UTC on the following two issues regarding my Tamil fractions/symbols proposal.

### §1. Unification of the sign for *nāli/paṭi* with OBF3 TAMIL DAY SIGN

Whenever a character is identical in shape but different in character properties (especially GC) in relation to an existing character or an also-proposed character, I have disunified them. These disunifications are listed in §5.1 of the proposal.

I have however not disunified one case when both the shape and character properties are the same but the meaning is different, since meanings are not a part of encoding. The case is that of the measure *nāli/paṭi* seen in §3.2 “Measures” of my proposal:

|    |                     |                  |                   |
|----|---------------------|------------------|-------------------|
| 6. | ௩                   | 3 uḷakku         | mūvuḷakku         |
| 7. | ௨ (ALREADY ENCODED) | 2 uri (4 uḷakku) | nāli / paṭi       |
| 8. | ௩                   | 8 nāli / paṭi    | kuruṇi / marakkāl |

This character is identical to OBF3 TAMIL DAY SIGN ௨ in both shape and character properties including GC=So. However, in collation, the symbol ௨ used as a measure would be sorted between ௩ and ௩ as seen above whereas that used as a day sign would be sorted with the other calendric items OBF5 ௩ TAMIL YEAR SIGN and OBF4 ௩ TAMIL MONTH SIGN.

I personally do not have experience with collation, but I think there can be no meaningful need for sorting measures along with calendrical signs in a single list, which is why I have not proposed to disunify based on meanings. I also have noted in §5.4 of the proposal that the same symbol is also used as an invocation at the beginning of texts, and certainly it would not be appropriate to yet again disunify the same written form for that!

I would like to have confirmation from the UTC that my choice to *not* disunify based on meaning is the correct approach. Advice on how the varied collation as per meaning could be handled would also be useful for potential implementers of collation.

## *§2. Numeric value for the proposed 11FD2 TAMIL FRACTION DOWNSCALING FACTOR KIIZH*

I have explained in §4.5 “Lower fractions” of my proposal that the proposed character  $\text{கீ}$  11FD2 TAMIL FRACTION DOWNSCALING FACTOR KIIZH when prefixed to fractions scales their value down by a factor of  $1/_{320}$  :

|                  |      |    |                                   |
|------------------|------|----|-----------------------------------|
| The value of     | சூ   | is | $3/4$                             |
| and the value of | கீசூ | is | $1/_{320} \times 3/4 = 3/_{1280}$ |
| The value of     | ௨    | is | $1/2$                             |
| and the value of | கீ௨  | is | $1/_{320} \times 1/2 = 1/_{640}$  |

... and so on.

I have allotted this character the numeric value property of  $1/320$ . To be precise, the line proposed for UnicodeData.txt is:

```
11FD2;TAMIL FRACTION DOWNSCALING FACTOR KIIZH;No;0;L;;;1/320;N;;;;;
```

The UTC should confirm that these properties are appropriate.

## *§3. Providing time to various parties for reviewing my proposal*

I also take this opportunity to place this matter on record:

Various parties interested in Tamil – especially the Gov’t of Tamil Nadu, ICTA of Sri Lanka, INFITT WG02 and perhaps others – would probably like to review my proposal and provide feedback. Therefore, while I request the UTC to take my proposal onto the agenda of the forthcoming August 2012 meeting for their own perusal and feedback, especially on the above two issues, I request that decisions as to encoding may be postponed to the November 2012 meeting to provide time for others to review and provide feedback. Hopefully interested parties will be timely in sending in their responses and appropriate revisions (if any) towards a finalized set of characters can be made by that time for a decision on encoding to be duly taken up by the UTC then.

-0-0-0-