உத்தமை
INFITT

உலகத் தமிழ்த் தகவல் தொழில்நுட்ப மன்றம்
International Forum for Information Technology in Tamil

முனைவர் வாசு ரெங்கநாதன்	முகுந்தராஜ் சுப்பிரமணியன்	மணி மு. மணிவண்ணன்
அமெரிக்கா	ஆஸ்திரேலியா	இந்தியா
தலைவர்	துணைத்தலைவர்	பணிக்குழு—2 (தமிழ் ஒருங்குறி) தலைவர்

To

Dr. Lisa Moore,
Chair, UNICODE Technical Committee
Unicode Consortium
U.S.A

Dear Dr. Lisa Moore,

Re: [**137-A50**] **Action item for** Mani Manivannan, Shriramana Sharma, Naga Ganesan: Work with experts to decide on a consistent approach to Tamil fraction names, by January 27, 2014 for the next UTC meeting.

It is our pleasure to inform you that after discussing the above subject with scholars in Tamil Nadu and within INFITT's Tamil Unicode Working Group WG-02 with members from around the world that includes both Dr. Naga Ganesan and Dr. Shriramana Sharma, we have arrived at a consensus as described in the enclosed document. These discussions considered all proposals including a modified list from Dr. Ganesan as well as the reduced form of Madras University Tamil Lexicon transliteration without diacritics by Dr. Sharma. Since there is no Tamil to Roman transcription standard and considering the limitations of using only upper case letters for character names, our choice was restricted to making these character names readable. The names adapt the convention normally found in telephone directories in Tamil Nadu where the Tamil names are transcribed to Roman letters.

Please let me know if you have any questions.

Regards,

மணி மு. மணிவண்ணன்

Mani M. Manivannan
Chair, INFITT Working Group 2 on Tamil Unicode

Dr. Vasu Renganathan (USA), Chairman	Dr. Kalyanasundaram (Switz.), Member	Mani M Manivannan, (Chair, India Chapter)
Mugunth Subramanian (Aus) – Vice-Chair	Iraama. Kiruttinan (India), Member	A. Elangovan (Treasurer, India Chapter)
Thavaruban (Sri Lanka) – Exec Director	Sasikumar Ranganathan (USA), Member	C. M. Elantamil (Chair, Malaysia Chapter)
Sornam Sankar (USA), Treasurer		

Spelling changes for Tamil fractions and symbols

By INFITT Working Group 2 on Tamil Unicode

Introduction

The proposal N4430 L2/13-047 (henceforth, "the AP" for "accepted proposal") to encode 55 new Tamil characters has been accepted by the UTC and WG2 committees at the meetings in 2013 May and June respectively. These characters are currently in the pipeline in Stage 5.

The names of some of the accepted characters involve Tamil words freely adapted as per popular convention (not formally transliterated) into the basic Latin script i.e. A to Z. The author of the AP has discussed the spelling of these names in the following documents:

- ∞ N4477 L2/13-175 – Response to L2/13-161 on naming Tamil fractions and symbols
- ∞ N4478 L2/13-176 – Request to change one character name from N4430 L2/13-047
- ∞ N4526 L2/14-018 – Spelling changes for Tamil fractions and symbols

Specifically, the last document above proposed a new model for adapting these words from Tamil in response to feedback from various parties, and, based on this model, requested to change the spelling of sixteen character names from the AP.

Changes required to proposed names

On discussion with various scholars and experts including the author of the above documents, it is found that it is sufficient to modify only three of the names w.r.t. the AP.

These changes are as follows:

1) For 11FD6 வது, the name TAMIL SIGN PADAKKU is proposed.

   Herein, the word PADAKKU (for பதக்கு) should be changed to PATHAKKU.

   The proposed annotation of 11FD7 ꩷ should also be changed accordingly.

2) For 11FDF கட, the name TAMIL SIGN GEJAM is proposed.

   Herein, the word GEJAM (for கெஜம்) should be changed to GAJAM (for கஜம்).

   This is to follow the primary form of the word as per N4478 L2/13-176.

   The proposed annotation of 11FE0 குட should also be changed accordingly.

3) For 11FE8 மி, the name TAMIL SIGN MUDALIYA is proposed.

   Herein, the word MUDALIYA (for முதலிய) should be changed to MUTHALIYA.

These are the only changes requested to the normative names w.r.t. the AP.

Changes required to existing/proposed annotations/aliases

The AP also requested certain editorial additions/modifications in terms of annotations and informative aliases. The spelling used in some of those requested additions/modifications also needs to be changed. Further, other changes not considered by the AP are required to existing annotations/aliases. These are as follows:

Changes for existing annotations/aliases not discussed by the AP

1) 0B83 ஃ currently has an alias AYTHAM for ஆய்தம்.

   This should be changed to AAYTHAM in this alias and in the following annotation.

2) 0BF6 ௶ currently has an alias PATRU for பற்று.

   This should be changed to PARRU.

3) 0BF9 ௹ currently has an alias RUPAI for ரூபாய்.

   This should be changed to RUUBAAY.

Changes for aliases proposed by the AP

Some aliases proposed by the AP for new characters need to be changed as follows:

		Tamil Word	Alias as per AP	Now requested
0BDF	நூ	நாளது	NAALADU	NAALATHU
0BFE	�P	போக	POKA	POGA
0BFF	ஆூ	ஆக	AAKA	AAGA
11FC0	ஹத	முந்திரி	MUNDIRI	MUNTHIRI
11FC3	௲	கால் வீசம்	KAALVIISAM	KAAL VIISAM
11FC5	சூ	அரை வீசம்	ARAIVIISAM	ARAI VIISAM
11FC7	கூ	முக்கால் வீசம்	MUKKAALVIISAM	MUKKAAL VIISAM
11FE7	ழூ	முதல்	MUDAL	MUTHAL

Further, for 0BF7 ௷, the AP proposes to replace the existing alias VARAVU by EDUPPU. However, while the new alias EDUPPU may be added, the existing alias VARAVU should not be removed since that word refers to a different meaning also denoted by the character.

All other additions/changes to annotations/aliases may be done as per the AP.

Summary of changes

Effectively, if the names and annotations/aliases are first input as requested by the AP and then the above changes are applied, the desired result would be achieved. For the convenience of the editors, we have attached herewith the UnicodeData.txt and NamesList.txt entries/diff which need to be added to the *present published standard* to achieve the desired spelling of names and annotations/aliases. These files include some more minor wording changes to the annotations required to glue in the main changes requested above so it is recommended to use these files directly.

Finally, it is important to record here the spelling model followed. To maximize readability, a free adaptation is followed rather than a strict transliteration which would anyway be difficult with the basic Latin script A-Z. The following mapping is used. Among multiple Latin letters for the same Tamil letter, the one best reflecting prevalent pronunciation is used as per de facto consensus among native users.

அ	ஆ	இ	ஈ	உ	ஊ	எ ஏ	ஐ	ஒ ஓ	ஒள
A	AA	I	II	U	UU	E	AI	O	AU

க	ச	ட	த	ப	ங ஞ ண ந ன	ம
K/G	C/J/S	T/D	TH	P/B	N	M

ய	ர ற	ல ள	வ	ழ
Y	R	L	V	ZH

ஸா ஷ	ஸ	ஜ	ஹ ஃ
SH	S	J	H

-o-o-o-

3