## Universal Multiple-Octet Coded Character Set

 International Organization for Standardization Organisation Internationale de NormalisationDoc Type: Working Group Document
Title: Comments on Lanna encoding in FPDAM4
Source: China
Status: Member Contribution
Action: For consideration

After carefully reading and discussing the Lanna encoding scheme in FPDAM4 and related documents, the Chinese experts in Yunnan province (where there lives over 400,00 Dai people) believe that the current scheme is uncompleted and is with errors or mistakes.

## 1. Script's name

The people which living in northern part of Thailand, Myanmar, Laos and Dai nation districts in China do not use this script as LANNA. The speakers of Northern Thai name this language "GAM MENG", means "local language", and name the script "TO MENG", means "local script", or "TO THAM", means "script of scripture". In China, people address this script "TO TAI", "TO TAILUE", "TO THAM", "script of scripture". Therefore, we regard that we should respect the people's will, using the name that people have identified.
2. Character naming:

Every alphabet has its own name in China's OLD TAILUE script project, called "TSUE TO". There are many homophone alphabets for OLD TAILUE script. Besides, different alphabets are used in different situations, and conjoin with different alphabets to express different meanings. On that account, alphabet names are for discriminating different alphabet, and making things convenient for every alphabet's use.

Most of consonant alphabets’ names are their pronunciation spelled with two vowels " O " and " A ", for instance, consonant k alphabet' name is "TO KO KA" and x called "TO XO XA" and so forth.

Except above alphabet name, alphabet has another name---"SING PHON",
when alphabet name is used to be spelled, it is hardly to spell directly due to too many syllables. In order to make the spelling easy, every alphabet has simple "SING PHON". "SING PHON" can be equal to alphabet pronunciation, just take one syllable and spell it with vowel.

In FPDAM4 LANNA(OLD TAILUE), "SING PHON" is used for most of alphabets’ names, but alphabets’ names themselves are used for some cases:

1A5A LANNA SIGN MAI KANG LAI
1A5B LANNA SIGN KHUEN MAI KANG LAI
1A74 LANNA SIGN MAI KANG
1A7B LANNA SIGN MAI SAM
1AA7 LANNA SIGN MAI YAMOK
Character names in FPDAM4 are different from those in NEW TAILUE, even if the same alphabets. We believe that since "NEW TAI LUE" is already encoded in ISO/IEC 10646, character names should be conform to those encoded, especially the same characters or the same phoneme. Characters with different names in FPDAM4 and NEW TAILUE are shown below:

Table 58 - Row 19: New Tai Lue


Plane 00

Table 62 - Row 1A: Lanna


The same characters with different names are as below:

| New Tai Lue | Lanna |
| :---: | ---: |
| 1980 | 1A4E |

1983 1A21
1986 1A24
1987 1A26
1988 1A27
1989 1A28
198B 1A2A
1990 1A56
1994 1A3A
199E 1A48
There are also many errors in the FPDAM4, especially the vowel alphabet names. Corrected vowel names are given below:

| 1A54 | VOWEL SIGN AE |
| :--- | :--- |
| 1A58 | component part HONG |
| 1A61 | VOWEL SIGN VOWEL SHORTENER |
| 1A65 | VOWEL SIGN AM |
| 1A66 | VOWEL SIGN AM |
| 1A67 | VOWEL SIGN I |
| 1A68 | VOWEL SIGN II |
| 1A69 | VOWEL SIGN UE |
| 1A6A | VOWEL SIGN UUE |
| 1A6B | VOWEL SIGN U |
| 1A6C | VOWEL SIGN UU |
| 1A6D | FINAL NG |
| 1A6E | VOWEL SIGN OA BELOW |
| 1A6F | VOWEL SIGN OY |
| 1A70 | VOWEL SIGN E |
| 1A71 | VOWEL SIGN AE |
| 1A72 | VOWEL SIGN OO |
| 1A73 | VOWEL SIGN AI |
| 1A75 | VOWEL SIGN OA ABOVE |
| 1A76 | VOWEL SIGN MAI KANG |
| 1A77 | SIGN TONE-1 |
| 1A78 | SIGN TONE-2 |

## 3. Ordering of characters

Characters 1A4E, 1A4F, 1A50, 1A51, 1A52, 1A53, 1A54, 1A55 are put after consonant alphabets in FPDAM4. Traditionally, these 8 glottal alphabets are put preceding all consonant alphabets, according to consonant's articulate segment attribution collate: glottal, velar, apical, labial. Thus, these 8 glottal should be ordered preceding all alphabets.

In FPDAM4, the same phoneme's characters are not arranged altogether. For instance, 6 phoneme " $S$ " characters 1A28, 1A29, 1A49, 1A4A, 1A4B, 1A5F are located at three positions respectively.

## 4. Characters should be deleted.

Some characters in FPDAM4 are the same characters in different handwriting, they are unlikely to appear in the same version, so it is unnecessary to be encoded. For example:

1A28 v.s. 1A4B: 1A28 is the different writing form of 1A4B, these two characters pronunciation and usage are totally the same, thus the encoding of 1A28 is not needed.

1 A 71 v.s. 12 A 7 : the same as that of 1 A 28 v.s. 1 A 4 B , thus 1 A 72 is not needed.

1 A 58 is a combining mark, it can combined with $k, t, K, ~ T, ~ U, ~ h, ~ s, ~ z, ~ a ~ . ~$ In FPDAM, 1A55 is a combination of 1A58 and 1A4E , so 1A55 is unnecessary to be encoded.

The combined character 1A5D has only one combining form in practice, i.e. 1A50. There is no need to encode 1A5D any more since 1A50 is encoded separately.
5. Characters should be added

The frequently used super joined characters 1A7C and 1A7D, which is proposed in N3207, are not in FPDAM4. 1A7C and 1A7D See characters table N3207:

TABLE XX - Row 1A: LANNA (Northern Thai font style)


There three characters need to be encoded: 1A5E, 1A60 1A61, see character table Old Tai Lue. The characters are named "TO TAI", means "dead character", can not be separated.

Our new character table：

TABLE XX－ROW 1A ：OLD TAILUE

| 0 | $\checkmark$ | $\bigcirc$ | $\infty$ | ¢ | \％ | of | $\bigcirc$ | $\bigcirc$ | $\oplus$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | ๕ | ब | ๑ | § | c） | $\bigcirc 0$ | $\bigcirc$ | $\bigcirc$ | 6 |
| 2 | ล | 03 | $\checkmark$ | $\delta$ | $\square$ | $\infty$ | $J$ | $J$ | 3 |
| 3 | વ్దీ | 003 | － | 6 | O | 80 | P | ¢ | \％ |
| 4 | 2 | Q | $\bigcirc$ | 13 | б́ | do | 9 | ¢ | © |
| 5 | 5 | ค | ¢ | 2 | Oo | O | Э | ${ }^{3}$ | \％ |
| 6 | e | $G$ | J | 3 | 0 | － | G | $\sigma$ | 3 |
| 7 | $\infty$ | ต | 0 | $\checkmark$ | $\bigcirc$ | ó | १ | $\downarrow$ | $?$ |
| 8 | a | ¢ | $\varepsilon$ | O | வ่ | $\delta$ | $\bigcirc$ | $\omega$ | 1 |
| 9 | § | $\omega$ | $\cdots$ | 8 | 8 | $\bigcirc$ | C | $\wedge$ | $\\|$ |
| A | $\bigcirc$ | 已 | $\infty$ | O＇ | 8 | $O^{\circ}$ | 雃 | 9 | 4 |
| B | 6 | $\omega$ | Q | $\delta^{8}$ | － | $\bigcirc$ |  | $\bigcirc$ | 4 |
| c | బు | g | ヵ | 8 | 8 | of | $y$ | \％ | $\infty$ |
| D | ค | 2 | $\bigcirc$ | 9 | ？ | 9 | 4 | x | 0 |
| E | $\bigcirc$ | బ | $\bigcirc$ | อె | $i_{i}$ |  | 3 | $5$ |  |
| F | 0 | $\infty$ | $\infty$ | $a$ | $\bigcirc$ |  |  | 多 |  |

$\mathrm{G}=00$
$\mathrm{P}=00$

## character names

| hex | Name |
| :---: | :---: |
| 1A20 | TAI LUE LETTER HIGH |
| 1A21 | TAI LUE LETTER LOW II |
| 1A22 | TAI LUE LETTER LOW UU |
| 1A23 | TAILUE LETTER LOW EE |
| 1A24 | TAI LUE LETTER HIGH XA |
| 1A25 | TAI LUE LETTER LOW XA |
| 1A26 | TAI LUE LETTER HIGH SA |
| 1A27 | TAI LUE LETTER HIGH SA |
| 1A28 | TAI LUE LETTER HIGH SA |
| 1A29 | TAI LUE LETTER HIGH SA |
| 1A2A | TAI LUE LETTER HIGH SA |
| 1A2B | TAILUE LETTER HIGH YA |
| 1A2C | TAILUE LETTER LOW YA |
| 1A2D | TAI LUE LETTER HIGH TA |
| 1A2E | TAI LUE LETTER HIGH THA |
| 1A2F | TAI LUE LETTER LOW THA |
| 1A30 | TAI LUE LETTER LOW NA |
| 1A31 | TAI LUE LETTER LOW NA |
| 1A32 | TAI LUE LETTER LOW PA |
| 1A33 | TAI LUE LETTER LOW HA |
| 1A34 | TAI LUE LETTER LOW LA |
| 1A35 | TAI LUE LETTER LOW LA |
| 1A36 | TAI LUE VOWEL SIGN AA |
| 1A37 | TAI LUE VOWEL SIGN I |
| 1A38 | TAI LUE VOWEL SIGN II |
| 1A39 | TAI LUE VOWEL SIGN UE |
| 1A3A | TAI LUE VOWEL SIGN UE |
| 1A3B | TAI LUE VOWEL SIGN UE |
| 1A3C | TAI LUE VOWEL SIGN U |
| 1A3D | TAI LUE VOWEL SIGN U |
| 1A3E | TAI LUE VOWEL SIGN U |
| 1A3F | TAI LUE VOWEL SIGN UU |
| 1A40 | TAI LUE VOWEL SIGN UU |
| 1A41 | TAI LUE VOWEL SIGN UU |
| 1A42 | TAI LUE VOWEL SIGN OA |
| 1A43 | TAI LUE VOWEL SIGN OA |
| 1A44 | TAI LUE VOWEL SIGN OA |
| 1A45 | TAI LUE VOWEL SIGN OOY |
| 1A46 | TAI LUE VOWEL SIGN OOY |


| hex | Name |
| :---: | :---: |
| 1A47 | TAI LUE VOWEL SIGN OOY |
| 1A48 | TAI LUE COMPONENT PART Y |
| 1A49 | TAI LUE COMPONENT PART Y |
| 1A4A | TAI LUE COMPONENT PART Y |
| 1A4B | TAI LUE COMPONENT PART P |
| 1A4C | TAI LUE COMPONENT PART TH |
| 1A4D | TAILUE COMPONENT PART P |
| 1A4E | TAI LUE COMPONENT PART V |
| 1A4F | TAI LUE LETTER FINAL NG |
| 1A50 | TAI LUE LETTER FINAL NG |
| 1A51 | TAI LUE LETTER FINAL NG |
| 1A52 | TAI LUE LETTER FINAL NG |
| 1A53 | TAI LUE LETTER FINAL N |
| 1A54 | TAI LUE LETTER FINAL N |
| 1A55 | TAI LUE LETTER FINAL N |
| 1A56 | TAI LUE LETTER FINAL N |
| 1A57 | TAI LUE LETTER FINAL N |
| 1A58 | TAI LUE LETTER FINAL N |
| 1A59 | TAI LUE LETTER FINAL N |
| 1A5A | TAI LUE LETTER FINAL N |
| 1A5B | TAI LUE LETTER FINAL N |
| 1A5C | TAI LUE LETTER FINAL M |
| 1A5D | TAI LUE LETTER FINAL M |
| 1A5E | TAI LUE LETTER FINAL M |
| 1A5F | TAI LUE LETTER FINAL K |
| 1A60 | TAI LUE LETTER FINAL K |
| 1A61 | TAI LUE LETTER FINAL K |
| 1A62 | TAI LUE LETTER FINAL K |
| 1A63 | TAI LUE LETTER FINAL T |
| 1A64 | TAI LUE LETTER FINAL T |
| 1A65 | TAI LUE LETTER FINAL T |
| 1A66 | TAI LUE LETTER FINAL T |
| 1A67 | TAI LUE LETTER FINAL T |
| 1A68 | TAI LUE LETTER FINAL T |
| 1A69 | TAI LUE LETTER FINAL T |
| 1A6A | TAI LUE LETTER FINAL T |
| 1A6B | TAI LUE LETTER FINAL T |
| 1A6C | TAI LUE LETTER FINAL T |
| 1A6D | TAI LUE LETTER FINAL T |


| hex | Name |
| :---: | :---: |
| 1A6E | TAI LUE LETTER FINAL T |
| 1A6F | TAI LUE LETTER FINAL T |
| 1AD0 | TAI LUE LETTER FINAL T |
| 1AD1 | TAI LUE LETTER FINAL T |
| 1AD2 | TAI LUE LETTER FINAL T |
| 1AD3 | TAI LUE LETTER FINAL T |
| 1AD4 | TAI LUE LETTER FINAL P |
| 1AD5 | TAI LUE LETTER FINAL P |
| 1AD6 | TAI LUE LETTER FINAL |
| 1AD7 | TAI LUE LETTER FINAL |
| 1AD8 | TAI LUE LETTER FINAL |
| 1AD9 | TAI LUE LETTER FINAL P |
| 1ADA | TAI LUE TONE MARK-1 |
| 1ADB | TAI LUE TONE MARK-2 |
| 1ADC | TAI LUE TONE MARK-2 |
| 1ADD | TAI LUE TONE MARK-2 |
| 1ADE | TAI LUE COMPONENT PART HONG |
| 1ADF | TAILUE WORD REPEAT SIGN |
| 1AE0 | TAILUE PUNCTUATION MARK |
| 1AE1 | TAILUE PUNCTUATION MARK |
| 1AE2 | TAILUE PUNCTUATION MARK |
| 1AE3 | TAI LUE DIGIT ONE |
| 1AE4 | TAI LUE DIGIT ONE |
| 1AE5 | TAI LUE DIGIT ONE |
| 1AE6 | TAI LUE DIGIT TWO |
| 1AE7 | TAI LUE DIGIT THREE |
| 1AE8 | TAI LUE DIGIT FOUR |
| 1AE9 | TAI LUE DIGIT FIVE |
| 1AEA | TAI LUE DIGIT SIX |
| 1AEB | TAI LUE DIGIT SEVEN |
| 1AEC | TAI LUE DIGIT EIGHT |
| 1AED | TAI LUE DIGIT NINE |
| 1AEE | TAILUEDIGIT NINE |

Dec

