2007-09-06

Universal Multiple-Octet Coded Character Set International Organization for Standardization Organisation Internationale de Normalisation

Doc 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
                        MAI KANG LAI
1A5B
     LANNA SIGN
                 KHUEN
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

| $\overline{}$ | Table 58 - Row | 13.1 | tew rai Luc |
|---------------|--|--------------|--|
| HEX | NAME | HEX | NAME |
| | | | |
| 1980 1981 | NEWTALUJE LETTERHICH QA NEWTALUJE LETTERLOW QA | 1900 190E | (This position shall not be used) NEW TATLUE SIGN LE |
| 1982 | NEW TAI LUE LETTER HIGH KA | 19DF | NEW TAILUE SIGN LEW |
| 1983 1984 | NEW TAI LUE LETTER HIGH XA NEW TAI LUE LETTER HIGH NGA | | |
| 1985 | NEW TAILUE LETTER HIGH NIGA NEW TAILUE LETTER LOW KA | 1 | |
| 1988 | NEW TAI LUE LETTER LOW XA | 1 | |
| 1987 1988 | NEW TAILUE LETTER LOWINGA NEW TAILUE LETTER HIGH TSA | 1 | |
| 1989 | NEW TAILUE LETTER HIGH SA | 1 | |
| 198A | NEW TAI LUE LETTER HIGH YA | 1 | |
| 198B 198C | NEW TAI LUE LETTER LOW TSA NEW TAI LUE LETTER LOW SA | 1 | |
| 1980 | NEW TAILUE LETTER LOW YA | 1 | |
| 198E | NEW TAILUE LETTER HIGH TA | 1 | |
| 198F 1990 | NEW TAILUE LETTER HIGH THA NEW TAILUE LETTER HIGH NA | 1 | |
| 1991 | NEW TAILUE LETTER LOW TA | 1 | |
| 1992 1993 | NEW TAILUE LETTER LOW THA | 1 | |
| 1993 | NEW TAILUE LETTER LOW NA NEW TAILUE LETTER HIGH PA | 1 | |
| 1995 | NEW TAI LUE LETTER HIGH PHA | 1 | |
| 1996 1997 | NEW TAILUE LETTER HIGH MA | 1 | |
| 1998 | NEW TAILUE LETTER LOW PA NEW TAILUE LETTER LOW PHA | 1 | |
| 1999 | NEW TAILUE LETTER LOW MA | 1 | |
| 199A 199B | NEWTALLUE LETTER HIGH FA NEWTALLUE LETTER HIGH VA | 1 | |
| 199C | NEW TAI LUE LETTER HIGH LA | 1 | |
| 1990 | NEW TAILUE LETTER LOW FA | 1 | |
| 199E 199F | NEW TAILUE LETTER LOW VA NEW TAILUE LETTER LOW LA | 1 | |
| 19AD | NEW TAI LUE LETTER HIGH HA | 1 | |
| 19A1 19A2 | NEW TAI LUE LETTER HIGH DA NEW TAI LUE LETTER HIGH BA | 1 | |
| 19A3 | NEW TAILUE LETTER LOWHA | 1 | |
| 19A4 | NEW TAILUE LETTER LOW DA | 1 | |
| 19A5 19A6 | NEW TAI LUE LETTER LOW BA NEW TAI LUE LETTER HIGH KVA | 1 | |
| 19A7 | NEW TAILUE LETTER HIGH XVA | 1 | |
| 1948 | NEW TAI LUE LETTER LOW KVA | 1 | |
| 19A9 19AA | NEW TAI LUE LETTER LOW XVA (This position shall not be used) | 1 | |
| 19AB | (This position shall not be used) | 1 | |
| 19AC | (This position shall not be used) (This position shall not be used) | 1 | |
| 19AD 19AE | (This position shall not be used) | 1 | |
| 19AF | (This position shall not be used) | 1 | |
| 1980 1981 | NEW TAI LUE VOWEL SIGN VOWEL SHORTENER NEW TAI LUE VOWEL SIGN AA | 1 | |
| 1982 | NEW TAI LUE VOWEL SIGN II | 1 | |
| 1983 | NEW TAILUE VOWEL SIGN U | 1 | |
| 1984 1985 | NEW TAI LUE VOWEL SIGN UU NEW TAI LUE VOWEL SIGN E | 1 | |
| 1986 | NEW TAI LUE VOWEL SIGN AE | | |
| 1997 1998 | NEW TAI LUE VOWEL SIGN O NEW TAI LUE VOWEL SIGN OA | 1 | |
| 1989 | NEW TAILUE VOWEL SIGN UE | 1 | |
| 198A | NEW TAI LUE VOWEL SIGN AY | 1 | |
| 1998 1980 | NEW TAI LUE VOWEL SIGN AAY NEW TAI LUE VOWEL SIGN UY | 1 | |
| 1980 | NEW TAILUE VOWEL SIGN OY | 1 | |
| 198E | NEW TAI LUE VOWEL SIGN CAY | 1 | |
| 198F 1900 | NEW TAI LUE VOWEL SION UEY NEW TAI LUE VOWEL SION IY | | |
| 1901 | NEW TAILUE LETTER FINAL V | | |
| 19C2 19C3 | MEW TAI LUE LETTER FINAL NG. MEW TAI LUE LETTER FINAL N | | |
| 1904 | NEW TAILUE LETTER FINAL M | | |
| 1905 | NEW TAILUE LETTER FINAL K | | |
| 19C8 19C7 | NEWTALLUE LETTERFINAL D NEWTALLUE LETTERFINAL B | | |
| 1908 | NEW TAILUE TONE MARK-1 | | |
| 1909 | NEW TAILUE TONE MARK-2 | | |
| 19CA 19CB | (This position shall not be used) (This position shall not be used) | | |
| 1900 | (This position shall not be used) | | |
| 1900 1908 | (This position shall not be used) (This position shall not be used) | | |
| 19CF | (This position shall not be used) | | |
| 1900 | NEW TAI LUE DIGIT ZERO | | |
| 19D1 19D2 | NEW TAILUE DIGIT ONE NEW TAILUE DIGIT TWO | | |
| 1903 | NEW TAILUE DIGIT THREE | | |
| 1904 | NEW TAI LUE DIGIT FOUR | | |
| 1905 1906 | NEW TAILUE DIGIT FIVE NEW TAILUE DIGIT SIX | | |
| 1907 | NEW TAI LUE DIGIT SEVEN | | |
| 1908 | NEW TAI LUE DIGIT EIGHT | | |
| 19D9 19DA | NEW TAI LUE DIGIT NINE (This position shall not be used) | | |
| 1908 | (This position shall not be used) | | |
| 19DC | (This position shall not be used) | | |
| ш | | \vdash | l |

Table 62 - Row 1A: Lanna

| HEX NAME 1A20 LANNA LETTER INGH KA 1921 LANNA LETTER INGH KAA 1922 LANNA LETTER INGH KOA 1A23 LANNA LETTER INGH KOA 1A24 LANNA LETTER INGH KOA 1A25 LANNA LETTER INGH KOA 1A25 LANNA LETTER INGH KOA 1A26 LANNA LETTER INGH KOA | 1A7D 1A7E | NAME (This position shall not be used) |
|---|---|--|
| 1021 ÚANNA LETTER HIGH KHA | | (This position shall not be used) |
| JAVAD LINNALETTER HIGH CA 1228 LINNALETTER HIGH CA 1229 LINNALETTER HIGH CA 1220 LINNALETTER HIGH CA 1220 LINNALETTER LOW CA 1220 LINNALETTER HIGH PATHAL 1220 LINNALETTER HIGH PATHAL 1220 LINNALETTER HIGH PATHAL 1220 LINNALETTER HIGH THAL 1220 LINNALETTER | 1/4/17 1/4/88 1/4/8 1/ | This postion shall not be used) LANNA DIGIT ONE LANNA DIGIT ONE LANNA DIGIT TWE LANNA DIGIT SEX LANNA DIGIT SEX LANNA DIGIT SEX LANNA DIGIT SEX LANNA DIGIT SEGURI LANNA DIGIT SEX LANNA DIGIT SEGURI THIS postion shall not be used) (This postion shall not seguri be used) (This postion shall not be used) |

Group 00 Plane 00 Row 1A

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
           VOWEL SIGN I
1A67
           VOWEL SIGN II
1A68
           VOWEL SIGN UE
1A69
1A6A
            VOWEL SIGN UUE
            VOWEL SIGN U
1A6B
            VOWEL SIGN UU
1A6C
1A6D
            FINAL NG
1A6E
            VOWEL SIGN OA BELOW
1A6F
            VOWEL SIGN OY
1A70
            VOWEL SIGN E
            VOWEL SIGN AE
1A71
1A72
            VOWEL SIGN OO
1A73
            VOWEL SIGN AI
            VOWEL SIGN OA ABOVE
1A75
            VOWEL SIGN MAI KANG
1A76
            SIGN TONE-1
1A77
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.

1A71 v.s. 12A7: the same as that of 1A28 v.s. 1A4B, thus 1A72 is not needed.

1A58 is a combining mark , it can combined with k_{ν} to K_{ν} To U_{ν} holds as z_{ν} 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)

| | 1A2 | 1A3 | 1A4 | 1A5 | 1A6 | 1A7 | 1A8 | 1A9 | 1AA | |
|---|-----|----------|------------|-------------|---------|-----------|-----|---------|------------|------------------|
| 0 | က | ಬ | ဘ | ಜ್ಞ | [.,.] | c | 0 | 0 | \oplus | |
| 1 | නු | ඔ | ಟ | වූ | ્દ | േ | 9 | a | | |
| 2 | (B) | 33 | ಬು | Ð | Ó | ि | J | 6 | ವೆ | |
| 3 | O | න | ಟ | (D) | ာ | ్ల | 9 | a) | •> | |
| 4 | ည) | വ | হ) | ع | ್ರೆ | ଥ | 9 | ß | @ | |
| 5 | ಖು | တ | ೯ | िल | ిం | ô · | ໆ | ß | 90% 90% | |
| 6 | 6 | 2 | ೲ | ස | ್ರಿ | <u></u> | G | ญี | 3 | |
| 7 | ۵ | <u> </u> | M 3 | 8 | ெ | Ö | ก | ಬ | ಬ್ | |
| 8 | သ | 0 | 0 | ি | ଣ | ò | O |) St. / | Ī | G = 00 P = 00 |
| 9 | ಇ | 0 | ග | ್ಟ | ್ರ | া | 6 | လ | TII. | |
| Α | ત | ಲ | ವಿತ್ರ | الماء الماء | 8 | ਂ | | | ণ | |
| В | 2 | ป | သ | ్ | ್ಫ | P. | | | 역 | |
| С | کر | ක් | ಬ | ව | ್ಷ | s) (_) | | | 20) | |
| D | 2 | F | 8 | 3 | | િલ | | | රී | |
| Е | ಬ | C | ಎ | g | <u></u> | | | | | |
| F | ච | ධ | ೭೩ | ဿ | ्र | ಂ | | | | |

15

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

| | 1A2 | 1A3 | 1A4 | 1A5 | 1A6 | 1A7 | 1A8 | 1A9 | 1AA | |
|---|------|-----|----------|------|-----|------------|----------------|------------|--------------------------|------------------|
| 0 | ಯ | చు | တ | a | హ | ୍ଧୀ | 0 | 0 | \oplus | |
| 1 | ಛ | G | ଭ | ક | လ် | c ਂ | 0 | 0 | Ф | |
| 2 | 53 | 23 | G | ย | | ೲ | J | J | యే | |
| 3 | 930 | ဿ | ۵ | 6 | ್ದಾ | So | 9 | ø | ♦ | |
| 4 | 9 | R | q | S | ó | 90 | 9 | သ | @ | |
| 5 | 9 | ವ | <u>е</u> | 2 | ಂ | † | 9 | ဌ | °00 €00 €00 €00 | |
| 6 | С | ଜ | J | 3 | ্ব | ċ | G | હ | 8 | |
| 7 | က | ន | O | ଧ | ೆ | Ċ | ૧ | ป | 2 | |
| 8 | а | ప | £ | ြ | া | ŏ | o | బ | 1 | G = 00 P = 00 |
| 9 | ક | ω | б | ્ | ి | ਂ | C | ٨ | ı | |
| Α | 0 | ည | ဘ | ाँ | ి | ಿ | - | ន្ន | ч | |
| В | 6 | ε | Q | ీ | ੇ | ं | | လ | 4 | |
| C | బు . | ဍ | 8 | 8 | ి | O. | 7/0 | | 82 | |
| D | ര | 2 | လ | ្វ | ု | · | 1 | 11 2014 | O | |
| E | Ð | బ | 0 | స్రొ | Q. | | odini Shiri | On the | | |
| F | သ | ന | S | ુ | ै | | | 100 | | |

character names

| hex | Name | | hex | Name |
|------|---------------------|-----|------|---------------------------|
| 1A20 | TAI LUE LETTER HIGH | I | 1A47 | TAI LUE VOWEL SIGN OOY |
| 1A21 | TAI LUE LETTER LOW | II | 1A48 | TAI LUE COMPONENT PART Y |
| 1A22 | TAI LUE LETTER LOW | UU | 1A49 | TAI LUE COMPONENT PART Y |
| 1A23 | TAILUE LETTER LOW | EE | 1A4A | TAI LUE COMPONENT PART Y |
| 1A24 | TAI LUE LETTER HIGH | XA | 1A4B | TAI LUE COMPONENT PART P |
| 1A25 | TAI LUE LETTER LOW | XA | 1A4C | TAI LUE COMPONENT PART TH |
| 1A26 | TAI LUE LETTER HIGH | SA | 1A4D | TAILUE COMPONENT PART P |
| 1A27 | TAI LUE LETTER HIGH | SA | 1A4E | TAI LUE COMPONENT PART V |
| 1A28 | TAI LUE LETTER HIGH | SA | 1A4F | TAI LUE LETTER FINAL NG |
| 1A29 | TAI LUE LETTER HIGH | SA | 1A50 | TAI LUE LETTER FINAL NG |
| 1A2A | TAI LUE LETTER HIGH | SA | 1A51 | TAI LUE LETTER FINAL NG |
| 1A2B | TAILUE LETTER HIGH | YA | 1A52 | TAI LUE LETTER FINAL NG |
| 1A2C | TAILUE LETTER LOW | YA | 1A53 | TAI LUE LETTER FINAL N |
| 1A2D | TAI LUE LETTER HIGH | TA | 1A54 | TAI LUE LETTER FINAL N |
| 1A2E | TAI LUE LETTER HIGH | THA | 1A55 | TAI LUE LETTER FINAL N |
| 1A2F | TAI LUE LETTER LOW | THA | 1A56 | TAI LUE LETTER FINAL N |
| 1A30 | TAI LUE LETTER LOW | NA | 1A57 | TAI LUE LETTER FINAL N |
| 1A31 | TAI LUE LETTER LOW | NA | 1A58 | TAI LUE LETTER FINAL N |
| 1A32 | TAI LUE LETTER LOW | PA | 1A59 | TAI LUE LETTER FINAL N |
| 1A33 | TAI LUE LETTER LOW | HA | 1A5A | TAI LUE LETTER FINAL N |
| 1A34 | TAI LUE LETTER LOW | LA | 1A5B | TAI LUE LETTER FINAL N |
| 1A35 | TAI LUE LETTER LOW | LA | 1A5C | TAI LUE LETTER FINAL M |
| 1A36 | TAI LUE VOWEL SIGN | | 1A5D | TAI LUE LETTER FINAL M |
| 1A37 | TAI LUE VOWEL SIGN | | 1A5E | TAI LUE LETTER FINAL M |
| 1A38 | TAI LUE VOWEL SIGN | | 1A5F | TAI LUE LETTER FINAL K |
| 1A39 | TAI LUE VOWEL SIGN | | 1A60 | TAI LUE LETTER FINAL K |
| 1A3A | TAI LUE VOWEL SIGN | | 1A61 | TAI LUE LETTER FINAL K |
| 1A3B | TAI LUE VOWEL SIGN | | | TAI LUE LETTER FINAL K |
| 1A3C | TAI LUE VOWEL SIGN | | 1A63 | TAI LUE LETTER FINAL T |
| 1A3D | TAILUE VOWEL SIGN | | 1A64 | TAI LUE LETTER FINAL T |
| 1A3E | TAI LUE VOWEL SIGN | | 1A65 | TAI LUE LETTER FINAL T |
| 1A3F | TAILUE VOWEL SIGN | | 1A66 | TAILUE LETTER FINAL T |
| 1A40 | TALLUE VOWEL SIGN | | 1A67 | TALLUE LETTER FINAL T |
| 1A41 | TALLUE VOWEL SIGN | | 1A68 | TALLUE LETTER FINAL T |
| 1A42 | TALLUE VOWEL SIGN | | 1A69 | TALLUE LETTER FINAL T |
| 1A43 | TAILLIE VOWEL SIGN | | 1A6A | TAILUE LETTER FINAL T |
| 1A44 | TALLUE VOWEL SIGN | | 1A6B | TALLUE LETTER FINAL T |
| 1A45 | TAILUE VOWEL SIGN | | 1A6C | TALLUE LETTER FINAL T |
| 1A46 | TAI LUE VOWEL SIGN | OUY | 1A6D | TAI LUE LETTER FINAL T |
| | | | | |

| hex | Name | Dec |
|------|-----------------------------|-----|
| 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 P | |
| 1AD7 | TAI LUE LETTER FINAL P | |
| 1AD8 | TAI LUE LETTER FINAL P | |
| 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 | Name |
|-----|------|
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |