ISO/IEC JTC1/SC2/WG2 N3313

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	
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

HEX	NAME	HEX	NAME
1980	NEW TAI LUE LETTER HIGH QA	19DD	(This position shall not be used)
1981	NEW TAI LUE LETTER LOW QA	19DE	NEW TAI LUE SIGN LE
1982	NEW TAI LUE LETTER HIGH KA	19DF	NEW TAI LUE SIGN LEW
1983	NEW TAI LUE LETTER HIGH XA		
1984	NEW TAI LUE LETTER HIGH NGA		
1985	NEW TAI LUE LETTER LOW KA		
1986	NEW TAI LUE LETTER LOW XA		
1987	NEW TAI LUE LETTER LOW NGA		
1988	NEW TAI LUE LETTER HIGH TSA		
1989	NEW TAI LUE LETTER HIGH SA		
198A	NEW TAI LUE LETTER HIGH YA		
198B	NEW TAI LUE LETTER LOW TSA		
198C	NEW TAI LUE LETTER LOW SA		
198D	NEW TAI LUE LETTER LOW YA		
198E	NEW TAI LUE LETTER HIGH TA		
198F	NEW TAI LUE LETTER HIGH THA		
1990	NEW TAI LUE LETTER HIGH NA		
1991	NEW TAI LUE LETTER LOW TA		
1992	NEW TAI LUE LETTER LOW THA		
1993	NEW TAI LUE LETTER LOW NA		
1994	NEW TAI LUE LETTER HIGH PA		
1995	NEW TAI LUE LETTER HIGH PHA		
1996	NEW TAI LUE LETTER HIGH MA		
1997	NEW TAI LUE LETTER LOW PA		
1998	NEW TAI LUE LETTER LOW PHA		
1999	NEW TAI LUE LETTER LOW MA		
199A	NEW TAI LUE LETTER HIGH FA		
199B	NEW TAI LUE LETTER HIGH VA		
199C	NEW TAI LUE LETTER HIGH LA		
199D	NEW TAI LUE LETTER LOW FA		
199E	NEW TAI LUE LETTER LOW VA		
199F	NEW TAI LUE LETTER LOW LA		
19A0	NEW TAI LUE LETTER HIGH HA		
19A1	NEW TAI LUE LETTER HIGH DA		
19A2	NEW TAI LUE LETTER HIGH BA		
19A3	NEW TAI LUE LETTER LOW HA		
19A4	NEW TAI LUE LETTER LOW DA		
19A5	NEW TAI LUE LETTER LOW BA		
19A6	NEW TAI LUE LETTER HIGH KVA		
19A7	NEW TAI LUE LETTER HIGH XVA		
19A8	NEW TAI LUE LETTER LOW KVA		
19A9	NEW TAI LUE LETTER LOW XVA		
19AA	(This position shall not be used)		
19AB	(This position shall not be used)		
19AC	(This position shall not be used)		
19AD	(This position shall not be used)		
19AE	(This position shall not be used)		
19AF	(This position shall not be used)		
19B0	NEW TAI LUE VOWEL SIGN VOWEL SHORTENER		
19B1	NEW TAI LUE VOWEL SIGN AA		
19B2	NEW TAI LUE VOWEL SIGN II		
19B3	NEW TAI LUE VOWEL SIGN U		
19B4	NEW TAI LUE VOWEL SIGN UU		
19B5	NEW TAI LUE VOWEL SIGN E		
19B6	NEW TAI LUE VOWEL SIGN AE		
19B7	NEW TAI LUE VOWEL SIGN O		
19B8	NEW TAI LUE VOWEL SIGN OA		
19B9	NEW TAI LUE VOWEL SIGN UE		
19BA	NEW TAI LUE VOWEL SIGN AY		
19BB	NEW TAI LUE VOWEL SIGN AAY		
19BC	NEW TAI LUE VOWEL SIGN UY		
19BD	NEW TAI LUE VOWEL SIGN OY		
19BE	NEW TAI LUE VOWEL SIGN OAY		
19BF	NEW TAI LUE VOWEL SIGN UEY		
19C0	NEW TAI LUE VOWEL SIGN IY		
19C1	NEW TAI LUE LETTER FINAL V		
19C2	NEW TAI LUE LETTER FINAL NG		
19C3	NEW TAI LUE LETTER FINAL N		
19C4	NEW TAI LUE LETTER FINAL M		
19C5	NEW TAI LUE LETTER FINAL K		
19C6	NEW TAI LUE LETTER FINAL D		
19C7	NEW TAI LUE LETTER FINAL B		
19C8	NEW TAI LUE TONE MARK-1		
19C9	NEW TAI LUE TONE MARK-2		
19CA	(This position shall not be used)		
19CB	(This position shall not be used)		
19CC	(This position shall not be used)		
19CD	(This position shall not be used)		
19CE	(This position shall not be used)		
19CF	(This position shall not be used)		
19D0	NEW TAI LUE DIGIT ZERO		
19D1	NEW TAI LUE DIGIT ONE		
19D2	NEW TAI LUE DIGIT TWO		
19D3	NEW TAI LUE DIGIT THREE		
19D4	NEW TAI LUE DIGIT FOUR		
19D5	NEW TAI LUE DIGIT FIVE		
19D6	NEW TAI LUE DIGIT SIX		
19D7	NEW TAI LUE DIGIT SEVEN		
19D8	NEW TAI LUE DIGIT EIGHT		
19D9	NEW TAI LUE DIGIT NINE		
19DA	(This position shall not be used)		
19DB	(This position shall not be used)		
19DC	(This position shall not be used)		

Table 62 - Row 1A: Lanna

HEX	NAME	HEX	NAME
1A20	LANNA LETTER HIGH KA	1A7D	(This position shall not be used)
1A21	LANNA LETTER HIGH KHA	1A7E	(This position shall not be used)
1A22	LANNA LETTER HIGH KXA	1A7F	LANNA COMBINING CRYPTOGRAMMIC DOT
1A23	LANNA LETTER LOW KA	1A80	LANNA DIGIT ZERO
1A24	LANNA LETTER LOW KXA	1A81	LANNA DIGIT ONE
1A25	LANNA LETTER LOW KHA	1A82	LANNA DIGIT TWO
1A26	LANNA LETTER NGA	1A83	LANNA DIGIT THREE
1A27	LANNA LETTER HIGH CA	1A84	LANNA DIGIT FOUR
1A28	LANNA LETTER HIGH CHA	1A85	LANNA DIGIT FIVE
1A29	LANNA LETTER KHUEN HIGH CHA	1A86	LANNA DIGIT SIX
1A2A	LANNA LETTER LOW CA	1A87	LANNA DIGIT SEVEN
1A2B	LANNA LETTER NORTHERN THAI LOW CA	1A88	LANNA DIGIT EIGHT
1A2C	LANNA LETTER LOW SA	1A89	LANNA DIGIT NINE
1A2D	LANNA LETTER NORTHERN THAI LOW SA	1A8A	(This position shall not be used)
1A2E	LANNA LETTER LOW CHA	1A8B	(This position shall not be used)
1A2F	LANNA LETTER NYA	1A8C	(This position shall not be used)
1A30	LANNA LETTER RATA	1A8D	(This position shall not be used)
1A31	LANNA LETTER HIGH RATHA	1A8E	(This position shall not be used)
1A32	LANNA LETTER DA	1A8F	(This position shall not be used)
1A33	LANNA LETTER LOW RATHA	1A90	LANNA THAM DIGIT ZERO
1A34	LANNA LETTER RANA	1A91	LANNA THAM DIGIT ONE
1A35	LANNA LETTER HIGH TA	1A92	LANNA THAM DIGIT TWO
1A36	LANNA LETTER HIGH THA	1A93	LANNA THAM DIGIT THREE
1A37	LANNA LETTER LOW TA	1A94	LANNA THAM DIGIT FOUR
1A38	LANNA LETTER LOW THA	1A95	LANNA THAM DIGIT FIVE
1A39	LANNA LETTER NA	1A96	LANNA THAM DIGIT SIX
1A3A	LANNA LETTER BA	1A97	LANNA THAM DIGIT SEVEN
1A3B	LANNA LETTER HIGH PA	1A98	LANNA THAM DIGIT EIGHT
1A3C	LANNA LETTER HIGH PHA	1A99	LANNA THAM DIGIT NINE
1A3D	LANNA LETTER HIGH FA	1A9A	(This position shall not be used)
1A3E	LANNA LETTER LOW PA	1A9B	(This position shall not be used)
1A3F	LANNA LETTER LOW FA	1A9C	(This position shall not be used)
1A40	LANNA LETTER LOW PHA	1A9D	(This position shall not be used)
1A41	LANNA LETTER MA	1A9E	(This position shall not be used)
1A42	LANNA LETTER LOW YA	1A9F	(This position shall not be used)
1A43	LANNA LETTER HIGH YA	1AA0	LANNA SIGN WIANG
1A44	LANNA LETTER RA	1AA1	LANNA SIGN WIANGWAAK
1A45	LANNA LETTER RUE	1AA2	LANNA SIGN SAWAN
1A46	LANNA LETTER LA	1AA3	LANNA SIGN KEOW
1A47	LANNA LETTER LUE	1AA4	LANNA SIGN HOY
1A48	LANNA LETTER WA	1AA5	LANNA SIGN DOKMAI
1A49	LANNA LETTER HIGH SHA	1AA6	LANNA SIGN REVERSED ROTATED RANA
1A4A	LANNA LETTER HIGH SSA	1AA7	LANNA SIGN MAI YAMOK
1A4B	LANNA LETTER HIGH SA	1AA8	LANNA SIGN KAAN
1A4C	LANNA LETTER HIGH HA	1AA9	LANNA SIGN KUANKUU
1A4D	LANNA LETTER LLA	1AAA	LANNA SIGN SATKAAN
1A4E	LANNA LETTER A	1AAB	LANNA SIGN SATKAANKUU
1A4F	LANNA LETTER NORTHERN THAI A	1AAC	LANNA SIGN HANG
1A50	LANNA LETTER I	1AAD	LANNA SIGN CAANG
1A51	LANNA LETTER II	1AAE	(This position shall not be used)
1A52	LANNA LETTER U	1AAF	(This position shall not be used)
1A53	LANNA LETTER UU		
1A54	LANNA LETTER EE		
1A55	LANNA LETTER OO		
1A56	LANNA LETTER LOW HA		
1A57	LANNA LETTER LAE		
1A58	LANNA CONSONANT SIGN MEDIAL RA		
1A59	LANNA CONSONANT SIGN MEDIAL LA		
1A5A	LANNA SIGN MAI KANG LAI		
1A5B	LANNA SIGN KHUEN MAI KANG LAI		
1A5C	LANNA CONSONANT SIGN FINAL NGA		
1A5D	LANNA CONSONANT SIGN LOW PA		
1A5E	LANNA CONSONANT SIGN HIGH RATHA OR LOW PA		
1A5F	LANNA LETTER GREAT SA		
1A60	LANNA SIGN SAKOT		
1A61	LANNA VOWEL SIGN A		
1A62	LANNA VOWEL SIGN MAI SAT		
1A63	LANNA VOWEL SIGN AA		
1A64	LANNA VOWEL SIGN TALL AA		
1A65	LANNA VOWEL SIGN I		
1A66	LANNA VOWEL SIGN II		
1A67	LANNA VOWEL SIGN UE		
1A68	LANNA VOWEL SIGN UUE		
1A69	LANNA VOWEL SIGN U		
1A6A	LANNA VOWEL SIGN UU		
1A6B	LANNA VOWEL SIGN O		
1A6C	LANNA VOWEL SIGN OA BELOW		
1A6D	LANNA VOWEL SIGN OY		
1A6E	LANNA VOWEL SIGN E		
1A6F	LANNA VOWEL SIGN AE		
1A70	LANNA VOWEL SIGN OO		
1A71	LANNA VOWEL SIGN AI		
1A72	LANNA VOWEL SIGN THAM AI		
1A73	LANNA VOWEL SIGN OA ABOVE		
1A74	LANNA SIGN MAI KANG		
1A75	LANNA SIGN TONE-1		
1A76	LANNA SIGN TONE-2		
1A77	LANNA SIGN KHUEN TONE-3		
1A78	LANNA SIGN KHUEN TONE-4		
1A79	LANNA SIGN KHUEN TONE-5		
1A7A	LANNA SIGN RA HAAM		
1A7B	LANNA SIGN MAI SAM		
1A7C	(This position 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
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.

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、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)

G = 00
P = 00

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

G = 00
P = 00

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 AA	1A5D	TAI LUE LETTER FINAL M
1A37	TAI LUE VOWEL SIGN I	1A5E	TAI LUE LETTER FINAL M
1A38	TAI LUE VOWEL SIGN II	1A5F	TAI LUE LETTER FINAL K
1A39	TAI LUE VOWEL SIGN UE	1A60	TAI LUE LETTER FINAL K
1A3A	TAI LUE VOWEL SIGN UE	1A61	TAI LUE LETTER FINAL K
1A3B	TAI LUE VOWEL SIGN UE	1A62	TAI LUE LETTER FINAL K
1A3C	TAI LUE VOWEL SIGN U	1A63	TAI LUE LETTER FINAL T
1A3D	TAI LUE VOWEL SIGN U	1A64	TAI LUE LETTER FINAL T
1A3E	TAI LUE VOWEL SIGN U	1A65	TAI LUE LETTER FINAL T
1A3F	TAI LUE VOWEL SIGN UU	1A66	TAI LUE LETTER FINAL T
1A40	TAI LUE VOWEL SIGN UU	1A67	TAI LUE LETTER FINAL T
1A41	TAI LUE VOWEL SIGN UU	1A68	TAI LUE LETTER FINAL T
1A42	TAI LUE VOWEL SIGN OA	1A69	TAI LUE LETTER FINAL T
1A43	TAI LUE VOWEL SIGN OA	1A6A	TAI LUE LETTER FINAL T
1A44	TAI LUE VOWEL SIGN OA	1A6B	TAI LUE LETTER FINAL T
1A45	TAI LUE VOWEL SIGN OOY	1A6C	TAI LUE LETTER FINAL T
1A46	TAI LUE VOWEL SIGN OOY	1A6D	TAI LUE LETTER FINAL T

hex	Name	Dec	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 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		