DATE: 2007-04-03

ISO/IEC JTC 1/SC 2/WG 2

Universal Multiple-Octet Coded Character Set (UCS) - ISO/IEC 10646

Secretariat: ANSI

DOC TYPE:	
TITLE:	Proposing on Encoding Old Tai Lue
SOURCE:	China
STATUS:	To be considered by WG2
DISTRIBUTION:	WG2 experts
MEDIUM:	Electronic
NO. OF PAGES:	
REFERENCE	WG2N3121, WG2N3161, WG2N3169, WG2N3207

A. Administrative

1. **Title:**	Proposing on Encoding Old Tai Lue
2. Requester's name:	China
3. Requester type (Member body/Liaison/Individual contribution):	Member body
4. Submission date:	2007-04-03
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	Yes
(or) More information will be provided later:	

B. Technical – General

1. Choose one of the following:	
a. This proposal is for a new script (set of characters):	Yes
Proposed name of script:	Old Tai Lue
b. The proposal is for addition of character(s) to an existing block:	No
Name of the existing block:	
2. Number of characters in proposal:	111
3. Proposed category (select one from below - see section 2.2 of P&P document):	A

A-Contemporary		B.1-Specialized (small collection)		B.2-Specialized (large collection)	
C-Major extinct		D-Attested extinct		E-Minor extinct	
F-Archaic Hieroglyphic or Ideographic				G-Obscure or questionable usage symbols	

4. Is a repertoire including character names provided?	Yes
a. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?	Yes
b. Are the character shapes attached in a legible form suitable for review?	Yes

5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? TrueType fonts from WEIFANG BEIDA JADE BIRD HUAGUANG TECHNOLOGY CO.,LTD.

If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used: Mr. Yin Jianmin jimyin@vip.sina.com or Mr. Chen Zhuang chenzh@cesi.ac.cn

6. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? yes

b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? yes

7. Special encoding issues:

Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? No

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see http://www.unicode.org/Public/UNIDATA/UCD.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? No

If YES explain

2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? Yes

If YES, with whom? Yunnan Nationality Publishing Company, Tai language group of Yunnan people's broadcasting station

If YES, available relevant documents:

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Yes

Reference: Old Tai Lue is used by Dai people in Yunnan province.

4. The context of use for the proposed characters (type of use; common or rare) common

Reference: It is widely used to write OLd Tai Lue language by Dai people.

5. Are the proposed characters in current use by the user community? Yes

If YES, where? Reference: Yunnan Province, China

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? Yes

If YES, is a rationale provided? Yes

If YES, reference: Contemporary use and accordance with the roadmap.

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	Yes
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	No
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	No
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?	No
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
11. Does the proposal include use of combining characters and/or use of composite sequences?	Yes
If YES, is a rationale for such use provided?	Yes
If YES, reference:	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	Yes
If YES, reference:	Details are in the proposal
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	No
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility character(s)?	No
If YES, is the equivalent corresponding unified ideographic character(s) identified?	
If YES, reference:	

Explanations of Old Tai Lue Script Encoding Proposal

Old Tai Lue script, also known as Xishuangbanna Old Tai Lue script, has been created since the spread of Hinayana and tightly related to Buddhism spread in communities of Dai ethnic group. Many people regard this script as "[to^{55}tham^{41}]", means script of scriptures, classic script. It is obvious that the purpose of creation is for the translation of Pali Buddhistic scriptures. This script is widely used in Xishuangbanna, and temples in Simao, Shuangjiang also use it for copy writing of scriptures. It has been used as traditional Buddhistic scriptures' script in northern part of Thailand, northern part of Myanmar, Laos and other countries and regions in Southeast Asia. Therefore, this script still possesses the position of classic. Scriptures of Dai people in Xishuangbanna are reputed as many as[pet^{35}mɯn^{35}si^{35}ban^{41}xan^{55}] , means 84 thousands volumes. Old Tai Lue scriptures, also called as palm leave scriptures, record not only Buddhistic teachings, but also history, literature, law, calendar, medicine and other aspects of Dai people's experiences and wisdoms in life. It is a valuable culture inheritance and treasure.

From the aspect of this script's usage, Old Tai Lue script has the longest history for using and widest coverage in Xishuangbanna. Dai people are all believe in Buddhism. Every man should be a Buddhist for a period in his life during which man receives the Buddhistic teachings and grasps the script. A great many of secularized monks bring this script to the folk, many traditional characters of Old Tai Lue script thus have been widely used in Dai people's daily life. In addition, Chinese government has positively contributed to the protection and inheritance of Old Tai Lue script. The script teaching in schools is a very effective way. Since 1954, bilingual teaching of Dai and Chinese has been adopted， text books have been translated and published, Tai Lue script has been taught from primary school. Since 1956, Old Tai Lue script weekly news paper has been published. Up to now, it has a history of 51years. In 1986, Autonomous government of Xishuangbanna promulgated a law to guarantee the status of Old Tai Lue script. Since then, Old Tai Lue script has been adopted in school education, news papers have also begun to use Old Tai Lue script. These activities greatly promote the development of Old Tai Lue script. In order to foster the capable person on Old Tai Lue script, Yunnan Nationalities University opens specialty to train talented people. Yunnan Nationalities Press has organized special group for editing and publishing readings of Tai Lue script since 1958. In 2002, *Chinese-Tai Lue Dictionary* had been published. At present, protection to palm scriptures written in Tai Lue script has been actively carried out. Over 4,000 palm scriptures have been collected in Xishuangbanna Autonomous Prefecture, and relevant protection and studies have been conducted. Since 2001, Xishuangbanna Autonomous government has invested in collecting, neatening, translating and publishing 100 volumes of *The collected edition of China Palm Scriptures*. Recently, 13 volumes has been published, 40 volumes have been translated. Old Tai Lue script is live in Dai people communities, and has a good mass base. Therefore, its historic value and practical significance are irreplaceable.

Old Tai Lue script belongs to phonetic script which is different from the improved New Tai Lue script. Old Tai Lue script is planar. A word can be in conjunction with vowels，codas, and tones which can be added at the right, left, overhead or under the character. That is to say there are marks at character's right, left, overhead and below.

This proposal collects 111 characters of Old Tai Lue script including consonants, vowels, under marks, overhead marks, tone marks and digits.

Encoding Ways

This proposal is under the principle of "select the different avoiding the same", and "encoding on the form" to determine the character's code. At the same time, encoded characters in New Tai Lue script proposal are not to be re-encoded.

Consonants encoding

This proposal specially encodes 22 consonant characters in Old Tai Lue script. These consonants can be used as base character in conjunction with other marks. Marks can be added at characters' front, back, overhead and beneath. Consonants can be separated in single and compound body. The single body is composed of independent element, and the compound body is composed of two or more elements. Independent encoding is adopted for the compound bodies. For example, the character 1A8B, 1A95 in the character table.

Labialized consonants: labialization is popular in Tai Lue script. Consonants subjoin with the labialized element [w] represented by the subjoined mark $\bar{O}$ under. Consonant is added at the overhead of the subjoined mark with above-under structure. Character is above and the mark is under. Labialized marks under are independently encoded in this proposal. See 1ACF in the character table.

Palatalized consonants: lingual exists in Old Tai Lue consonants. Consonants subjoin with the palatalized element [j] reminding the palatalization. Its representing form is subjoined with palatalized mark, see 1AA9 in the character table.

Consonants with the mark "[-": A great many of characters in Old Tai Lue script are called [$to^{55}ju^{35}hoŋ^{41}$], represented by the combination with the mark"[-" or 1ADE in the table. There are two situations after the subjoined consonants: 1. from unaspirated to aspirated when pronounced; 2. subjoined to be slap compound consonants, this mark shows the slap [l] of compound consonants.

Vowels encoding

In Old Tai Lue script, 18 are monophthongs, 13 diphthongs with [i] coda and [u] coda, 60 monophthongs with consonant codas.

As to monophthongs, except the encoded characters in New Tai Lue script proposal, this proposal has encoded 13 monophthongs which had been abrogated in New Tai Lue script proposal.

As to diphthongs, there are 13 diphthongs, in which 7 combine with [i] codas, 6 with [u] codas. Except the diphthong [ai], [a:i] with [i] coda have independent marks ᦵ 、ᧀ , the other diphthongs mostly are formed by vowels combining [i] coda, see the character 1AA9 in character table. Diphthongs with [u] coda are formed by vowels combining with mark *Ō*.

A great proportion of vowels in Old Tai Lue scripts are formed by monophthongs combining consonant codas, 60 altogether. There are 6 consonant codas, [n], [ŋ], [m], [k], [t], [p] respectively. Each of them has a different mark. In this proposal, all codas have been encoded independently.

Tone marks and their encoding

There are 9 tones represented by different syllables and two tone marks. No marks for the Tone-1, Tone-2 and Tone-3 have tone marks. According to tone marks used in Xishuangbanna, 4 tone marks are encoded in this proposal, tow marks for Tone-2, and another two for Tone-3. The 4 marks are positioned above the characters.

Digital marks

Based on the international principal of collecting all that can be seen, 15 digital marks have been encoded, digit 1 have three represented ways and digit 9 have two represented ways. Digital characters are independently encoded.

The structure of characters formed with the combination of consonant and vowel in Old Tai Lue script

Many structures can be formed in the combination of consonant and vowel in Old Tai Lue script. Furthermore, due to the long time unstandardization of the Old Tai Lue scripts, it is difficult to introduce all forms of the character structures. This paper can only briefly

introduce several structures. On the other hand, all tone marks are positioned overhead, introduction to character structures ignore tone marks.

1. k j **consonant + vowel parallel structure, vowel positioned behind consonant, suitable to all monophthongs positioned behind consonants.**

2. e k **vowel + consonant parallel structure, vowel positioned in front of the consonant, suitable to all vowels positioned in front of the consonants.**

3. e k A **vowel + consonant + short vowel mark parallel structure, a short vowel form with the vowel positioned in front of consonant, suitable to all monophthongs positioned in front of consonants.**

4. ◌◌◌g **vowel + consonant + coda parallel structure, vowel positioned in front of the consonant and the coda behind, suitable to all monophthongs with consonant coda positioned in front of consonants.**

5. t i **consonant + vowel vertical structure, vowel positioned above consonant, suitable to all monophthongs positioned above consonants.**

6. l ◌ **consonant + vowel vertical structure, vowel positioned under the consonant, suitable to all monophthongs positioned under consonants.**

7. h_r **consonant + vowel + coda parallel structure, vowel positioned behind the consonant and coda behind the vowel, suitable to all monophthongs with consonant coda positioned behind consonants.**

8. e k◌j **vowel + consonant + vowel mark + vowel vertical structure, vowel positioned in front of or back of consonant, and vowel mark above. Monophthong [ə] and diphthongs [əi] and [au] usually appear in this kind of structure.**

9. k◌A **consonant + vowel + vowel vertical structure, short vowel form with monophthongs positioned above consonant, suitable to all monophthongs above consonants.**

10. V k| **vowel + consonant + vowel vertical structure, vowel positioned in front of and under consonant, mainly the representing form of vowel [**O**].**

11. e l| **vowel + consonant + vowel vertical structure, vowel positioned in front of and above consonant, mainly the representing form of vowel [**e**].**

12. z j| **consonant + vowel + coda vertical structure, vowel positioned behind the consonant and coda under the vowel, suitable to vowel [a] with all consonant codas.**

13. V k' **consonant + vowel + coda vertical structure, vowel positioned in front of and under consonant, mainly the representing form of vowel [**O**] with consonant coda [**n**],[**ŋ**],[**m**] and [t].**

14. h|r **consonant + vowel + coda vertical structure, vowel positioned above consonant and consonant coda behind consonant, mainly the representing form of vowel [i] and vowel [ɯ] with consonant coda.**

15. e k| **vowel + consonant + vowel + coda vertical structure, vowel positioned in front of and above consonant, coda under consonant, representing form of monophthong [ə] with consonant coda.**

16. e h|d **vowel + consonant + vowel + coda vertical structure, vowel positioned in front of and above consonant, coda behind consonant, a combination of vowel [ə] of consonant coda with compound letters, labialized consonant, palatalized consonant.**

17. e k|A **vowel + consonant + vowel + vowel + vowel mark vertical structure, vowel positioned in front of, above, under, and behind consonant, representing short vowel form of monophthong [ə] .**

TABLE XX – Row 1A: OLD TAI LUE

	1A8	1A9	1AA	1AB	1AC	1AD	1AE
0	ᦨ	ᦩ	ᦪ	ᦫ	᦬	᦭	᦮
1	᦯	ᦰ	ᦱ	ᦲ	ᦳ	ᦴ	ᦵ
2	ᦶ	ᦷ	ᦸ	ᦹ	ᦺ	ᦻ	ᦼ
3	ᦽ	ᦾ	ᦿ	ᧀ	ᧁ	ᧂ	ᧃ
4	ᧄ	ᧅ	ᧆ	ᧇ	ᧈ	ᧉ	᧊
5	᧋	᧌	᧍	᧎	᧏	᧐	᧑
6	᧒	᧓	᧔	᧕	᧖	᧗	᧘
7	᧙	᧚	᧛	᧜	᧝	᧞	᧟
8	᧠	᧡	᧢	᧣	᧤	᧥	᧦
9	᧧	᧨	᧩	᧪	᧫	᧬	᧭
A	᧮	᧯	᧰	᧱	᧲	᧳	᧴
B	᧵	᧶	᧷	᧸	᧹	᧺	᧻
C	᧼	᧽	᧾	᧿	ᨀ	ᨁ	ᨂ
D	ᨃ	ᨄ	ᨅ	ᨆ	ᨇ	ᨈ	ᨉ
E	ᨊ	ᨋ	ᨌ	ᨍ	ᨎ	ᨏ	ᨐ
F	ᨑ	ᨒ	ᨓ	ᨔ	ᨕ	ᨖ	

character names

hex	Name
1A80	TAI LUE LETTER HIGH I
1A81	TAI LUE LETTER LOW II
1A82	TAI LUE LETTER LOW UU
1A83	TAILUE LETTER LOW EE
1A84	TAI LUE LETTER HIGH XA
1A85	TAI LUE LETTER LOW XA
1A86	TAI LUE LETTER HIGH SA
1A87	TAI LUE LETTER HIGH SA
1A88	TAI LUE LETTER HIGH SA
1A89	TAI LUE LETTER HIGH SA
1A8A	TAI LUE LETTER HIGH SA
1A8B	TAILUE LETTER HIGH YA
1A8C	TAILUE LETTER LOW YA
1A8D	TAI LUE LETTER HIGH TA
1A8E	TAI LUE LETTER HIGH THA
1A8F	TAI LUE LETTER LOW THA
1A90	TAI LUE LETTER LOW NA
1A91	TAI LUE LETTER LOW NA
1A92	TAI LUE LETTER LOW PA
1A93	TAI LUE LETTER LOW HA
1A94	TAI LUE LETTER LOW LA
1A95	TAI LUE LETTER LOW LA
1A96	TAI LUE VOWEL SIGN AA
1A97	TAI LUE VOWEL SIGN I
1A98	TAI LUE VOWEL SIGN II
1A99	TAI LUE VOWEL SIGN UE
1A9A	TAI LUE VOWEL SIGN UE
1A9B	TAI LUE VOWEL SIGN UE
1A9C	TAI LUE VOWEL SIGN U
1A9D	TAI LUE VOWEL SIGN U
1A9E	TAI LUE VOWEL SIGN U
1A9F	TAI LUE VOWEL SIGN UU
1AA0	TAI LUE VOWEL SIGN UU
1AA1	TAI LUE VOWEL SIGN UU
1AA2	TAI LUE VOWEL SIGN OA
1AA3	TAI LUE VOWEL SIGN OA
1AA4	TAI LUE VOWEL SIGN OA
1AA5	TAI LUE VOWEL SIGN OOY
1AA6	TAI LUE VOWEL SIGN OOY

hex	Name
1AA7	TAI LUE VOWEL SIGN OOY
1AA8	TAI LUE COMPONENT PART Y
1AA9	TAI LUE COMPONENT PART Y
1AAA	TAI LUE COMPONENT PART Y
1AAB	TAI LUE COMPONENT PART P
1AAC	TAI LUE COMPONENT PART TH
1AAD	TAILUE COMPONENT PART P
1AAE	TAI LUE COMPONENT PART V
1AAF	TAI LUE LETTER FINAL NG
1AB0	TAI LUE LETTER FINAL NG
1AB1	TAI LUE LETTER FINAL NG
1AB2	TAI LUE LETTER FINAL NG
1AB3	TAI LUE LETTER FINAL N
1AB4	TAI LUE LETTER FINAL N
1AB5	TAI LUE LETTER FINAL N
1AB6	TAI LUE LETTER FINAL N
1AB7	TAI LUE LETTER FINAL N
1AB8	TAI LUE LETTER FINAL N
1AB9	TAI LUE LETTER FINAL N
1ABA	TAI LUE LETTER FINAL N
1ABB	TAI LUE LETTER FINAL N
1ABC	TAI LUE LETTER FINAL M
1ABD	TAI LUE LETTER FINAL M
1ABE	TAI LUE LETTER FINAL M
1ABF	TAI LUE LETTER FINAL K
1AC0	TAI LUE LETTER FINAL K
1AC1	TAI LUE LETTER FINAL K
1AC2	TAI LUE LETTER FINAL K
1AC3	TAI LUE LETTER FINAL T
1AC4	TAI LUE LETTER FINAL T
1AC5	TAI LUE LETTER FINAL T
1AC6	TAI LUE LETTER FINAL T
1AC7	TAI LUE LETTER FINAL T
1AC8	TAI LUE LETTER FINAL T
1AC9	TAI LUE LETTER FINAL T
1ACA	TAI LUE LETTER FINAL T
1ACB	TAI LUE LETTER FINAL T
1ACC	TAI LUE LETTER FINAL T
1ACD	TAI LUE LETTER FINAL T

hex	Name
1ACE	TAI LUE LETTER FINAL T
1ACF	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