

Universal Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная Организация по Стандартизации

Doc Type: Working Group Document
Title: Proposal to encode the Tai Xaaü script in UCS
Source: China National Body
Status: National Body Contribution
Action: For consideration by ISO/IEC JTC1/SC2/WG2
Date: 2025-04-22

1 Introduction

The Tai Xaaü (, tai⁴³xa:u³³, *tai xaaü*) script is an Abugida script written by the Dai people (傣族) of Yunnan Province (云南省), PRC, to record the Jinping dialect of the Dai language (傣语金平方言). The total population of the Dai people in China is over 1.32 million (2021 statistics), mainly distributed in Yunnan Province and a few in Sichuan Province (四川省).

The Dai people have a long history. Like the Zhuang (壮族), Buyei (布依族), Dong (侗族), etc., who are close in origin, the Dai people are also descendants of the ancient Baiyue (百越), of which the Dai people are part of the Dianyue (滇越). According to the records of the Dianyue in *Shiji* (史记) and *Hanshu* (汉书), during the Han Dynasty (汉朝), the ancestors of the Dai people lived mainly in present-day south-western Yunnan Province.

The Dai people can be divided into several branches according to their endonyms, such as Tai Lue (傈僳) for those living in Xishuangbanna (西双版纳), Tai Taue (傣德) and Tai Le (傣哪) for those living in Dehong (德宏), Gengma (耿马), Menglian (孟连) and so on, Tai Xaaü (傣皓) for those living in Jinping (金平) county of Honghe (红河), and so on. These endonyms have the same first syllable and different second syllable, inferring that they originate from the same ethnic group. After the establishment of PRC, they are collectively known as the Dai people.

The Dai language (傣语) belongs to the Zhuang–Dai (or Tai) languages of the Zhuang–Dong (or Dong–Tai) languages of the Sino–Tibetan language family, with more than one million speakers in Yunnan Province and Sichuan Province, PRC. According to the difference in phonology and vocabulary, there are four dialects of the Dai language in China: Dehong dialect (德宏方言, ISO 639–3: [tdd](#)), Xishuangbanna dialect (西双版纳方言, ISO 639–3: [khh](#)), Honghe–Jinsha dialect (红金方言, ISO 639–3: [tiz](#)) and Jinping dialect (金平方言, ISO 639–3: [twh](#)).

The Tai Xaaü script is written by the Tai Xaaü people, mainly in Jinping Miao, Yao, and Dai Autonomous County (金平苗族瑶族傣族自治县), Honghe Hani and Yi Autonomous Prefecture (红河哈尼族彝族自治州), and a few in Jiangcheng Hani and Yi Autonomous County (江城哈尼族彝族自治县), Pu'er City (普洱市), both in Yunnan Province (云南省), PRC, to record the Jinping dialect of the Dai language, so it is also known as Jinping Dai writing system (金平傣文).

The Tai Xaa script is an Abugida script in which consonant letters behave as base units with inherited vowels, while vowel letters behave as additional units and cannot appear alone. Similar to Tai Tham and New Tai Lue, the Tai Xaa script contains two sets of consonant letters, constructed as oppositions of voiceless and voiced consonants in Proto-Tai, and expressed as higher and lower tone classes. However, unlike Tai Tham and New Tai Lue, the Tai Xaa script has no subjoined consonants, that is, all consonant letters behave identically.

2 Script details

2.1 Relations with Tai Viet

The Tai Xaa script is related to the Tai Viet script (ISO 15924: Tavt), which is encoded in ISO/IEC 10646:2003/Amd.5:2007 (and in Unicode Standard Version 5.2), as it can be noted that some letters are similar in their glyphs and sounds. According to L2/06-041 (Ngô & Brase, 2006) and L2/07-039R (Brase, 2007), the Tai Viet script was said to be constructed from several scripts, namely Black Tai, White Tai, Red Tai and Lao Song, to form a universal script for use in Vietnam, Laos and Thailand. The result of the Tai Viet character set, “is focused on the use of the script by the Tai Dam of Sơn La Province, Vietnam”, “contains the traditional Sơn La character set as a whole, plus three pairs of aspirated consonants (note: high/low kho, cho and pho) required by the Tai Dón language of Lai Châu Province”. L2/08-217 (Brase, 2008) further analyzed the Tai Don writing systems and concluded that the 13 extra characters should be disunified and encoded, and mentioned that, “Other readers may wish to argue for the disunification of other glyphs, or for the disunification of the entire Tai Don/Jinping script from Tai Viet. All such contributions and advice are earnestly sought”.

From the point of view of China’s needs, we agree with the latter opinion, that is, to disunify Tai Xaa (the so-called Jinping script) from Tai Viet. The main reasons are:

- The whole style of the Tai Viet representative glyphs presented in the [Code Charts](#) is unacceptable to the Chinese Tai Xaa community, and it makes users think that Tai Viet and Tai Xaa are two completely different scripts.
- When consider the fonts like *Noto Sans Tai Viet* and *Tai Heritage*, even for the similar glyphs, Tai Xaa and Tai Viet will have completely different character names for them. First, Tai Xaa and Tai Viet have opposite definitions for high consonants and low consonants, and it is unacceptable to name a high consonant as a low consonant. Second, some letters are considered to be used to record different phonemes in Tai Viet and in Tai Xaa (for example,  HIGH PO in Tai Viet corresponds to  LOW MO in Tai Xaa,  LOW PHO in Tai Viet corresponds to  HIGH FO in Tai Xaa), which will introduce confusion.

We are also aware of the font implementation issues that may arise after merging Tai Viet and Tai Xaa. On the one hand, there is currently no solution that is acceptable to both China and Vietnam. On the other hand, since Tai Viet was encoded early, the users have expressed concerns about the foreseeable lack of localization of the common characters after merging Tai Viet and Tai Xaa, no matter how similar their shapes may be.

In addition, some brief research has been done on the issue of the relationship between the four writing systems covered by Tai Viet. We have considered such materials as [Origin Myth in Tai Dam](#), [multiple writing systems for Tai Don](#), [untitled manuscript in Tai Daeng](#) and [the Song Petburi font for Thai Song](#). We agree with the preliminary view in [L2/06-041 \(Ngô & Brase, 2006\)](#) that the internal differences between the four are smaller than the differences between them and the other already encoded scripts, and their encoding can therefore be done as a subsequent extension of Tai Viet.

We also notice that the Tai Viet and Tai Xaaü correspond more or less closely to scripts used in neighboring regions, such as Tai Tham, Tai Le, Thai, Lao, *etc.*, due to their common Brāhmī origin. In conclusion, we suggest disunifying Tai Viet and Tai Xaaü, because of the great differences from the representative glyphs to the character names.

2.2 Encoding model

The existence of pre-base vowel letters in the Tai Xaaü script, as in Tai Tham, New Tai Lue and so on, makes it possible to consider different encoding models for the Tai Xaaü script, that is, those based on visual order and those based on logical order. After several meetings and discussions, this document suggests a visual order-based encoding model.

The main reason for choosing the visual order is that all consonants in the Tai Xaaü script behave identically as the base characters. Many Brāhmī model-based scripts are rich in consonant conjuncts, but the Tai Xaaü script is not. This leads to the fact that even when a control character is introduced as a vowel killer, it does not play any role and introduces confusion instead.

In addition, as will be mentioned later, the vowel letters and tone marks used as combining marks are placed more flexible, as they are only there to ensure that the whole syllable is distinguishable from the others. Using visual order makes it simpler to support the many possible positions of these combining marks.

2.3 Character names

The Latinization of the Tai Xaaü phonology has been determined for use in the character names, as shown below.

/b/	/d/	/f/	/h/	/k/	/k ^h /	/l/	/m/	/n/	/ŋ/	/ɲ/
B	D	F	H	K	KH	L	M	N	NG	NY
/p/	/p ^h /	/ʔ/	/s/	/t/	/t ^h /	/ts̃/	/ts̃ ^h /	/v/	/x/	/j/
P	PH	Q	S	T	TH	TS	TSH	V	X	Y
/k ^w /	/x ^w /	/ŋ ^w /	/h ^w /	/k ^{hw} /	/ʔ ^w /					
KW	XW	NGW	HW	KHW	QW					
/a/	/a:/	/i/	/e/	/ɛ/	/ɔ/	/o/	/u/	/ɯ/	/ə/	
A	AA	I	E	AE	O	OA	U	UE	EI	

A and AA. For closed syllables, the phoneme /a/ has long and short oppositions. In some documents it is written as /a/ and /a:/, in others as /ă/ and /a/, and in this document,

they are transcribed as A and AA. For open syllables, the phoneme /a/ has no long and short oppositions and is transcribed as A.

2.4 Consonant letters

There are 44 consonant letters:

๑ HIGH QO	๑๕ HIGH HO YO	๑๗ HIGH HO MO	๑๗ HIGH HO VO
๑ LOW QO	๑๖ LOW YO	๑๗ LOW MO	๑๖ LOW VO
๗ HIGH KO	๗ HIGH TO	๕ HIGH FO	๑๕ HIGH KHO
๗ LOW KO	๗ LOW TO	๑๖ LOW FO	๑๖ LOW KHO
๑๕ HIGH XO	๑๕ HIGH THO	๑๖ HIGH BO	๑๕ HIGH TSHO
๑๕ LOW XO	๑๖ LOW THO	๑๗ LOW BO	๑๕ LOW TSHO
๑๖ HIGH HO NGO	๑๗ HIGH HO NO	๑๕ HIGH HO LO	๑๗ HIGH HO NYO
๑ LOW NGO	๑๖ LOW NO	๑๖ LOW LO	๑๖ LOW NYO
๑๕ HIGH TSO	๑๕ HIGH PO	๑๗ HIGH HO	
๑๖ LOW TSO	๑๖ LOW PO	๑๕ LOW HO	
๑๕ HIGH SO	๑๖ HIGH PHO	๑๕ HIGH DO	
๑๖ LOW SO	๑๖ LOW PHO	๑๕ LOW DO	

High consonant letters with “ho”. There are seven high consonant letters with “ho” in their character names, because the characters are named according to how people pronounce these letters. For example, ๑ HIGH QO is pronounced /to³³ʔo³³/, ๑ LOW QO is pronounced /to³³ʔo⁴³/, ๗ HIGH KO is pronounced /to³³kʔo³³/, ๗ LOW KO is pronounced /to³³kʔo⁴³/, and so on. And for the seven letters, for example, ๑๖ HIGH HO NGO is pronounced as /hʔo³³ŋʔo³³/, ๑๕ HIGH HO YO is pronounced as /hʔo³³jʔo³³/, and so on. /to³³/ refers to a classifier, and therefore does not appear in character names. /hʔo³³/ indicates that the first part of these letters comes from the letter ๑ HIGH HO.

Labialization indicator. The consonant letter ๑ LOW VO is also used for indicating labialized consonant. For example, in the word ๑๗๑๖ (kwaen) “loosen”, ๑๗๑๖ is the consonant cluster for /k^w/ (if the inherited vowel is not considered). For a Brāhmī model, this consonant cluster would need to be encoded as <๗ HIGH KO, virama, ๑ LOW VO>, but it is indistinguishable from <๗ HIGH KO, ๑ LOW VO>. People pronounce this consonant cluster as /k^wa³³/ (kwa), because when people pronounce it with /ɔ/ (o), they find that the labialization is not reflected (/k^wɔ³³/ ~ /kʔo³³/), so they switch to using /a/.

High Qo. The consonant letter ๑ HIGH QO is also used as a vowel letter for /ɔ/. For example, in the word ๑๑๑๖ (pod) “lung”, ๑ HIGH QO is used for /ɔ/.

Consonants for coda. There are eight consonant letters that are used as a coda in syllables, they are: ๑ LOW YO, ๑ LOW VO, ๑ LOW NGO, ๑ LOW NO, ๑ LOW MO, ๗ HIGH KO, ๑๕ HIGH DO, and ๑ HIGH BO. They are used in syllables that end with /i/ (or /j/), /u/ (or /w/), /ŋ/, /n/, /m/, /k/, /t/, and /p/.

Initials. There are two types of initials in Tai Xaa, those using non-labialized single consonants, which use the 44 consonants mentioned above, and those using labialized single consonants, which do not appear in the modern texts but do appear in pre-modern literature. Eleven consonants (high/low ko, xo, ngo, ho and kho, plus high qo) have been identified that can be followed by a labialization mark, but it should be assumed that all consonants with a labialization mark constitute legal strings.

2.5 Vowel letters and vowel marks

There are 7 vowel letters and 11 vowel marks (presented with dotted circle):

• MAAI KA-1	∘ MAAI KI	∘ MAAI KU-1	• MAAI KAUE
∘ MAAI KA-2	∘ MAAI KE	∘ MAAI KU-2	• MAAI KAU
∘ MAAI KA-3	• MAAI KAE	∘ MAAI KU-3	• MAAI KAI
∘ MAAI KA-4	∘ MAAI TSAM	∘ MAAI KUE	
∘ MAAI KA-5	• MAAI KOA	∘ MAAI KEI	

Maai Ka. There are one vowel letter and four vowel marks that are pronounced as *maai ka*, and they can be grouped into two categories. • MAAI KA-1 and ∘ MAAI KA-2 can be used in open syllables. Note that the vowel mark ∘ MAAI KA-2 can only be used in open syllables with tone 3, and the vowel letter • MAAI KA-1 can also be used in closed syllables for long /a/. The other three vowel marks ∘ MAAI KA-3, ∘ MAAI KA-4, and ∘ MAAI KA-5 can only be used in closed syllables. Since there is no opposition of long and short /a/ in open syllables in Jinping dialect, they are all pronounced *maai ka*.

The three vowel marks ∘ MAAI KA-3, ∘ MAAI KA-4, and ∘ MAAI KA-5 are used with specific and complementary coda letters. ∘ MAAI KA-3 is combined with • LOW NO or • HIGH DO to form • /an/ or • /at/. ∘ MAAI KA-4 is combined with • LOW NGO or • HIGH KO to form • /aŋ/ or • /ak/. ∘ MAAI KA-5 is combined with • LOW MO or • HIGH BO to form • /am/ or • /ap/. Note that the two vowel marks ∘ MAAI KA-4 and ∘ MAAI KA-5 are rather flexible, as they can be positioned above the consonant letters in initial and in final. It should be assumed that whatever consonant letter the two vowel marks are placed on, the character sequences corresponding to the syllables they form are equivalent. For example, the character sequences • /lak⁵⁵/ is equivalent to • /lak⁵⁵/.

Maai Kae and Maai Kau. The pre-base vowel letter • MAAI KAE is used for vowel /ɛ/, as in • /kɛu³³/, and the pre-base and post-base vowel letter • MAAI KAU is used for vowels /au/, as in • /ʔau³³/ . Because the glyphs of these two characters are very similar, it should be assumed that the character sequence of two • MAAI KAU is confusable to one • MAAI KAE.

Maai Ku. There are one vowel letter and two vowel marks that are pronounced as *maai ku*. ∘ MAAI KU-1 can be used in any syllables. • MAAI KU-2 can only be used with • LOW NGO, for example, in • /luŋ⁴³/ . ∘ MAAI KU-3 can only be used with • LOW MO, for example, in • /tsum³³/ .

Finals. All the finals in the modern text are listed in Fig. 2, which includes 9 finals with monophthongs, 14 finals with diphthongs, and 62 finals with consonant coda. Further finals appear in pre-modern literature, e.g. the single letter • MAAI KU-2 and the single letter ∘ MAAI KA-5 are also used as finals.

2.6 Tone marks

There are 2 tone marks:

• MAI SAK	• MAI HOI
-----------	-----------

There are eight tone classes in Jinping dialect, according to [Dao et al. \(2019\)](#):

	A	B	C	D
high	1 / ³³ /	2 / ⁵⁵ /	3 / ²⁵ /	7 / ⁵⁵ /
low	4 / ⁴³ /	5 / ⁵² /	6 / ³¹ /	8 / ⁴³ /

For the eight tone classes:

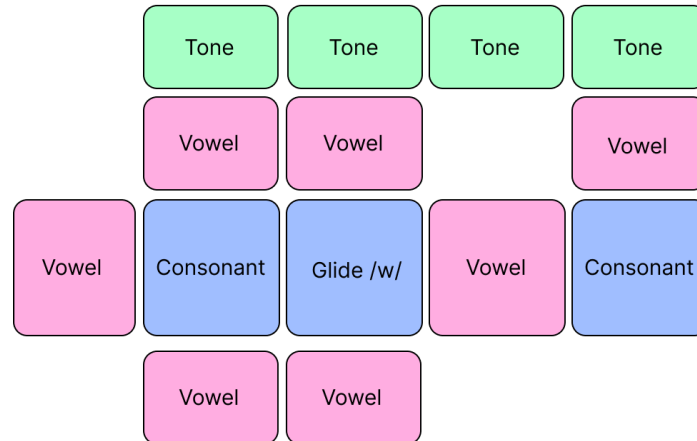
n , /ka ³³ /	n HIGH KO, , MAAI KA-1
h , /ka ⁵⁵ /	n HIGH KO, o MAI SAK, , MAAI KA-1
h , /ka ²⁵ /	n HIGH KO, o MAI HOI, , MAAI KA-1
h , /ka ⁴³ /	n LOW KO, , MAAI KA-1
h , /ka ⁵² /	n LOW KO, o MAI SAK, , MAAI KA-1
h , /ka ³¹ /	n LOW KO, o MAI HOI, , MAAI KA-1
n h /kak ⁵⁵ /	n HIGH KO, n HIGH KO, o MAAI KA-4
n h /kak ⁴³ /	n LOW KO, n HIGH KO, o MAAI KA-4

As can be seen above, **o** MAI SAK is used for tone 2 and tone 5, and **o** MAI HOI is used for tone 3 and tone 6.

In the Tai XaaU script, the tone is used to modify the whole syllable, so its position is relatively flexible, and in the literature, it can be added above any grapheme cluster. Therefore, it should be assumed that the resulting character sequence is equivalent regardless of the grapheme cluster to which the tone mark is added. For example, the character sequence **h****h****h** /sin⁵⁵/ is equivalent to **h****h****h**.

2.7 Syllable structure

The structure of each syllable in the Tai XaaU script can be summarized as:



The structure of a Tai XaaU syllable can be described as **OW?VC?T?**, where:

- **O** is a required onset, and it can be any of 44 consonant letters in section 2.4;
- **W** is an optional labialization indicator **o** LOW VO;
- **V** is a required vowel letter or vowel mark, where **,** MAAI KA-1 and **h** MAAI KOA are post-base vowel letters, **o** MAAI KAU should be used twice (one pre-base and one post-base), the other four vowel letters are pre-base;
- **C** is an optional coda, and it can be any of the eight consonant letters as described in section 2.4; the **C** and **V** will form a ligature **h** or **h**, as described in section 2.5;
- **T** is an optional tone mark, and it can be any of two as described in section 2.6.

For examples:

word	input
ໂອ /ʔau ³³ / “want”	ໂ MAAI KAU (V), ອ HIGH QO (O), ໂ MAAI KAU (V)
ໂນ /kau ²⁵ / “nine”	ໂ MAAI KAU (V), ນ HIGH KO (O), ອ MAI HOI (T), ໂ MAAI KAU (V)
ໂນ /pau ⁵² / “who”	ໂ MAAI KAUE (V), ນ LOW PHO (O), ອ MAI SAK (T)
ໂນ /neu ³³ / “sticky”	ໂ HIGH HO NO (O), ອ MAAI KE (V), ອ LOW VO (C)
ໂນ /sɔŋ ³³ / “two”	ໂ HIGH SO (O), ອ HIGH QO (V), ອ LOW NGO (C)
ໂນ /mak ⁴³ / “love”	ໂ LOW MO (O), ນ HIGH KO (C), ອ MAAI KA-4 (V)
ໂນ /mwi ⁵⁵ / “drizzle”	ໂ HIGH HO MO (O), ອ MAAI KUE (V), ອ MAI SAK (T), ອ LOW YO (C)
ໂນ /huŋ ⁵² / “daybreak”	ໂ MAAI KU-2 (V), ອ LOW HO (O), ອ MAI SAK (T), ອ LOW NGO (C)
ໂນ /x ^w a ³³ / “right”	ໂ HIGH XO (O), ອ LOW VO (W), ອ MAAI KA-1 (V)
ໂນ /x ^w aŋ ³³ / “horizontal”	ໂ HIGH XO (O), ອ LOW VO (W), ອ MAAI KA-1 (V), ອ LOW NGO (C)
ໂນ /k ^w ɛŋ ³³ / “soup”	ໂ MAAI KAE (V), ນ HIGH KO (O), ອ LOW VO (W), ອ LOW NGO (C)
ໂນ /k ^w in ³³ / “eat”	ໂ HIGH KO (O), ອ LOW VO (W), ອ MAAI KI (V), ອ LOW NO (C)
ໂນ /k ^{hw} an ³³ / “scarf”	ໂ HIGH KHO (O), ອ LOW VO (W), ອ LOW NO (C), ອ MAAI KA-3 (V)

2.8 Symbols and punctuation

There are 5 symbols with specific meanings:

ໂ NUENG “one”	ໂ BEN “moon”
ໂ KUN “person”	ໂ MAN “he”
ໂ MA “horse”	

Note that one should use ນ MAN and should not use <ໂ LOW MO, ອ MAAI KA-3>.

There are 7 Tai Xaaü-specific punctuation that are used for indicating the end of a sentence or an article:

ໂ MAAI SUT	ໂ LAN KAAŋ-1
ໂ MAAI SAM	ໂ LAN KAAŋ-2
ໂ MAAI SAM KOAK	ໂ LAN KAAŋ-3
ໂ MAAI SAM MAET	

The rest of the punctuation is identical to that used in Han ideographs.

2.9 Collation

The syllables in the Tai Xaaü script are ordered by initials, finals and tones.

- The order of initials is from ອ HIGH QO to ນ LOW NYO, then from <ໂ HIGH QO, ອ LOW VO> to <ໂ LOW NYO, ອ LOW VO>.
- The order of finals is (where “-” refers to an initial):
a (ໂ, ອ), i (ໂ), e (ໂ), ae (ໂ), o (ໂ), oa (ໂ), u (ໂ), ue (ໂ), ei (ໂ);
ai (ໂ), aai (ໂ), oi (ໂ), oai (ໂ), ui (ໂ), uei (ໂ), eii (ໂ);
au (ໂ), aau (ໂ), iu (ໂ), eu (ໂ), aeu (ໂ), eiu (ໂ);

aue (ʔ);
 ang (ᵛ), aang (ᵛᵛ), ing (ᵛᵛ), eng (ᵛᵛ), aeng (u-ᵛ), ong (ᵛᵛ), oang (ᵛᵛ),
 ung (ᵛᵛ < ᵛᵛ), ueng (ᵛᵛ), eing (ᵛᵛ);
 an (ᵛᵛ), aan (ᵛᵛ), ..., ein (ᵛᵛ);
 am (ᵛᵛ), aam (ᵛᵛ), ..., um (ᵛᵛ < ᵛᵛ), ..., eim (ᵛᵛ);
 ak (ᵛᵛ), aak (ᵛᵛ), ..., eik (ᵛᵛ);
 at (ᵛᵛ), aat (ᵛᵛ), ..., eit (ᵛᵛ);
 ap (ᵛᵛ), aap (ᵛᵛ), ..., eip (ᵛᵛ).

- The order of tones is: no tone, ̂ MAI SAK, ̂ MAI HOI.

Note that, the syllables recorded in the Tai Xaaui script are basically the original form of the morphemes, but in spoken language, the same morphemes change when they form different words, depending on their contexts, resulting in inconsistencies between the reading forms and the written forms. For examples, when /maʔ⁵⁵/ “gain” appears in the first syllable of a compound word, it should be read as /m⁵⁵/, thus changing its final; or, when a syllable in tone 4 follows a syllable in tone 3, the tone value /⁴³/ will change to /⁴⁵/, as /kun⁴³/ “person” in /ha²⁵kun⁴⁵/ “five persons”.

The order of Tai Xaaui characters is:

ᵛ HIGH QO < ᵛ LOW QO < ᵛ HIGH KO < ᵛ LOW KO < ᵛ HIGH XO < ᵛ LOW XO <
 ᵛᵛ HIGH HO NGO < ᵛᵛ LOW NGO < ᵛᵛ HIGH TSO < ᵛᵛ LOW TSO < ᵛᵛ HIGH SO <
 ᵛᵛ LOW SO < ᵛᵛ HIGH HO YO < ᵛᵛ LOW YO < ᵛᵛ HIGH TO < ᵛᵛ LOW TO <
 ᵛᵛ HIGH THO < ᵛᵛ LOW THO < ᵛᵛ HIGH HO NO < ᵛᵛ LOW NO < ᵛᵛ HIGH PO <
 ᵛᵛ LOW PO < ᵛᵛ HIGH PHO < ᵛᵛ LOW PHO < ᵛᵛ HIGH HO MO < ᵛᵛ LOW MO <
 ᵛᵛ HIGH FO < ᵛᵛ LOW FO < ᵛᵛ HIGH BO < ᵛᵛ LOW BO < ᵛᵛ HIGH HO LO <
 ᵛᵛ LOW LO < ᵛᵛ HIGH HO < ᵛᵛ LOW HO < ᵛᵛ HIGH DO < ᵛᵛ LOW DO <
 ᵛᵛ HIGH HO VO < ᵛᵛ LOW VO < ᵛᵛ HIGH KHO < ᵛᵛ LOW KHO < ᵛᵛ HIGH TSHO <
 ᵛᵛ LOW TSHO < ᵛᵛ HIGH HO NYO < ᵛᵛ LOW NYO <
 ̂ MAI SAK < ̂ MAI HOI <
 ̂ MAAI KA-1 < ̂ MAAI KA-2 < ̂ MAAI KA-3 < ̂ MAAI KA-4 < ̂ MAAI KA-5 <
 ̂ MAAI KI < ̂ MAAI KAE < ̂ MAAI KE < ̂ MAAI TSAM < ̂ MAAI KOA <
 ̂ MAAI KU-1 < ̂ MAAI KU-2 < ̂ MAAI KU-3 < ̂ MAAI KUE < ̂ MAAI KEI <
 ̂ MAAI KAUE < ̂ MAAI KAU < ̂ MAAI KAI <
 ̂ NUENG < ̂ KUN < ̂ MA < ̂ BEN < ̂ MAN <
 ̂ MAAI SUT < ̂ MAAI SAM < ̂ MAAI SAM KOAK < ̂ MAAI SAM MAET <
 ̂ LAN KAA-1 < ̂ LAN KAA-2 < ̂ LAN KAA-3

3 Proposed characters

3.1 Repertoire

The proposed repertoire is shown as below.

Codepoint	Glyph	Character Name
U+XX00		TAI XAAU LETTER HIGH QO
U+XX01		TAI XAAU LETTER LOW QO
U+XX02		TAI XAAU LETTER HIGH KO
U+XX03		TAI XAAU LETTER LOW KO
U+XX04		TAI XAAU LETTER HIGH XO
U+XX05		TAI XAAU LETTER LOW XO
U+XX06		TAI XAAU LETTER HIGH HO NGO
U+XX07		TAI XAAU LETTER LOW NGO
U+XX08		TAI XAAU LETTER HIGH TSO
U+XX09		TAI XAAU LETTER LOW TSO
U+XX0A		TAI XAAU LETTER HIGH SO
U+XX0B		TAI XAAU LETTER LOW SO
U+XX0C		TAI XAAU LETTER HIGH HO YO
U+XX0D		TAI XAAU LETTER LOW YO
U+XX0E		TAI XAAU LETTER HIGH TO
U+XX0F		TAI XAAU LETTER LOW TO
U+XX10		TAI XAAU LETTER HIGH THO
U+XX11		TAI XAAU LETTER LOW THO

U+XX12		TAI XAAU LETTER HIGH HO NO
U+XX13		TAI XAAU LETTER LOW NO
U+XX14		TAI XAAU LETTER HIGH PO
U+XX15		TAI XAAU LETTER LOW PO
U+XX16		TAI XAAU LETTER HIGH PHO
U+XX17		TAI XAAU LETTER LOW PHO
U+XX18		TAI XAAU LETTER HIGH HO MO
U+XX19		TAI XAAU LETTER LOW MO
U+XX1A		TAI XAAU LETTER HIGH FO
U+XX1B		TAI XAAU LETTER LOW FO
U+XX1C		TAI XAAU LETTER HIGH BO
U+XX1D		TAI XAAU LETTER LOW BO
U+XX1E		TAI XAAU LETTER HIGH HO LO
U+XX1F		TAI XAAU LETTER LOW LO
U+XX20		TAI XAAU LETTER HIGH HO
U+XX21		TAI XAAU LETTER LOW HO
U+XX22		TAI XAAU LETTER HIGH DO
U+XX23		TAI XAAU LETTER LOW DO
U+XX24		TAI XAAU LETTER HIGH HO VO
U+XX25		TAI XAAU LETTER LOW VO
U+XX26		TAI XAAU LETTER HIGH KHO
U+XX27		TAI XAAU LETTER LOW KHO

U+XX28		TAI XAAU LETTER HIGH TSHO
U+XX29		TAI XAAU LETTER LOW TSHO
U+XX2A		TAI XAAU LETTER HIGH HO NYO
U+XX2B		TAI XAAU LETTER LOW NYO
U+XX30		TAI XAAU MAAI KA-1
U+XX31		TAI XAAU MAAI KA-2
U+XX32		TAI XAAU MAAI KA-3
U+XX33		TAI XAAU MAAI KA-4
U+XX34		TAI XAAU MAAI KA-5
U+XX35		TAI XAAU MAAI KI
U+XX36		TAI XAAU MAAI KE
U+XX37		TAI XAAU MAAI KAE
U+XX38		TAI XAAU MAAI TSAM
U+XX39		TAI XAAU MAAI KOA
U+XX3A		TAI XAAU MAAI KU-1
U+XX3B		TAI XAAU MAAI KU-2
U+XX3C		TAI XAAU MAAI KU-3
U+XX3D		TAI XAAU MAAI KUE
U+XX3E		TAI XAAU MAAI KEI
U+XX3F		TAI XAAU MAAI KAUE
U+XX40		TAI XAAU MAAI KAU
U+XX41		TAI XAAU MAAI KAI

U+XX42	◌ [◌]	TAI XAAU TONE MAI SAK
U+XX43	◌ [◌]	TAI XAAU TONE MAI HOI
U+XX44	◌ [◌]	TAI XAAU SYMBOL NUENG
U+XX45	◌ [◌]	TAI XAAU SYMBOL KUN
U+XX46	◌ [◌]	TAI XAAU SYMBOL MA
U+XX47	◌ [◌]	TAI XAAU SYMBOL BEN
U+XX48	◌ [◌]	TAI XAAU SYMBOL MAN
U+XX49	◌ [◌]	TAI XAAU SYMBOL MAAI SUT
U+XX4A	◌ [◌]	TAI XAAU SYMBOL MAAI SAM
U+XX4B	◌ [◌]	TAI XAAU SYMBOL MAAI SAM KOAK
U+XX4C	◌ [◌]	TAI XAAU SYMBOL MAAI SAM MAET
U+XX4D	◌ [◌]	TAI XAAU SYMBOL LAN KAAAN-1
U+XX4E	◌ [◌]	TAI XAAU SYMBOL LAN KAAAN-2
U+XX4F	◌ [◌]	TAI XAAU SYMBOL LAN KAAAN-3

3.2 Character properties

The information in `NamesList.txt` is as follows.

```

@@      XX00      Tai Xaau      XX4F
@       Consonants
XX00    TAI XAAU LETTER HIGH QO
        x (tai viet letter low o - AA8E)
XX01    TAI XAAU LETTER LOW QO
        x (tai viet letter high o - AA8F)
XX02    TAI XAAU LETTER HIGH KO
        x (tai viet letter low ko - AA90)
XX03    TAI XAAU LETTER LOW KO
XX04    TAI XAAU LETTER HIGH XO
XX05    TAI XAAU LETTER LOW XO
XX06    TAI XAAU LETTER HIGH HO NGO
XX07    TAI XAAU LETTER LOW NGO
        x (tai viet letter high ngo - AA91)
XX08    TAI XAAU LETTER HIGH TSO
        x (tai viet letter low co - AA92)
XX09    TAI XAAU LETTER LOW TSO
        x (tai viet letter high co - AA93)

```

XX0A	TAI XAAU LETTER HIGH SO
XX0B	TAI XAAU LETTER LOW SO
XX0C	TAI XAAU LETTER HIGH HO YO
XX0D	TAI XAAU LETTER LOW YO
	x (tai viet letter high yo - AAA5)
XX0E	TAI XAAU LETTER HIGH TO
	x (tai viet letter low to - AA94)
XX0F	TAI XAAU LETTER LOW TO
	x (tai viet letter high to - AA95)
XX10	TAI XAAU LETTER HIGH THO
	x (tai viet letter low tho - AA96)
XX11	TAI XAAU LETTER LOW THO
XX12	TAI XAAU LETTER HIGH HO NO
	x (tai viet letter low no - AA98)
XX13	TAI XAAU LETTER LOW NO
	x (tai viet letter high no - AA99)
XX14	TAI XAAU LETTER HIGH PO
XX15	TAI XAAU LETTER LOW PO
XX16	TAI XAAU LETTER HIGH PHO
XX17	TAI XAAU LETTER LOW PHO
XX18	TAI XAAU LETTER HIGH HO MO
XX19	TAI XAAU LETTER LOW MO
XX1A	TAI XAAU LETTER HIGH FO
XX1B	TAI XAAU LETTER LOW FO
XX1C	TAI XAAU LETTER HIGH BO
	x (tai viet letter low bo - AA9A)
XX1D	TAI XAAU LETTER LOW BO
	x (tai viet letter high bo - AA9B)
XX1E	TAI XAAU LETTER HIGH HO LO
	x (tai viet letter low lo - AAA8)
XX1F	TAI XAAU LETTER LOW LO
XX20	TAI XAAU LETTER HIGH HO
	x (tai viet letter low ho - AAAC)
XX21	TAI XAAU LETTER LOW HO
	x (tai viet letter high ho - AAAD)
XX22	TAI XAAU LETTER HIGH DO
XX23	TAI XAAU LETTER LOW DO
	x (tai viet letter high do - AA93)
XX24	TAI XAAU LETTER HIGH HO VO
	x (tai viet letter low vo - AAAA)
XX25	TAI XAAU LETTER LOW VO
	x (tai viet letter high vo - AAAB)
XX26	TAI XAAU LETTER HIGH KHO
XX27	TAI XAAU LETTER LOW KHO
	x (tai viet letter high kho - AA83)
XX28	TAI XAAU LETTER HIGH TSHO
	x (tai viet letter low cho - AA8C)
XX29	TAI XAAU LETTER LOW TSHO
	x (tai viet letter high cho - AA8D)
XX2A	TAI XAAU LETTER HIGH HO NYO
XX2B	TAI XAAU LETTER LOW NYO
@	Vowels
XX30	TAI XAAU MAAI KA-1
	x (tai viet vowel aa - AAB1)
XX31	TAI XAAU MAAI KA-2
XX32	TAI XAAU MAAI KA-3
XX33	TAI XAAU MAAI KA-4
	x (tai viet mai kang - AAB0)
XX34	TAI XAAU MAAI KA-5
	x (tai viet vowel am - AABE)
XX35	TAI XAAU MAAI KI
	x (tai viet vowel i - AAB2)
XX36	TAI XAAU MAAI KE

x (tai viet vowel ia - AAB8)
 XX37 TAI XAAU MAAI KAE
 * precedes consonant in visual order
 x (tai viet vowel e - AAB5)
 XX38 TAI XAAU MAAI TSAM
 XX39 TAI XAAU MAAI KOA
 x (tai viet vowel ua - AABA)
 XX3A TAI XAAU MAAI KU-1
 x (tai viet vowel u - AAB4)
 XX3B TAI XAAU MAAI KU-2
 * precedes consonant in visual order
 x (tai viet vowel o - AAB6)
 XX3C TAI XAAU MAAI KU-3
 XX3D TAI XAAU MAAI KUE
 x (tai viet vowel ue - AAB3)
 XX3E TAI XAAU MAAI KEI
 x (tai viet mai khit - AAB7)
 XX3F TAI XAAU MAAI KAUE
 * precedes consonant in visual order
 x (tai viet vowel aue - AAB8)
 XX40 TAI XAAU MAAI KAU
 * precedes and follows consonant in visual order
 x (tai viet vowel uea - AAB9)
 XX41 TAI XAAU MAAI KAI
 * precedes consonant in visual order
 x (tai viet vowel ay - AABC)
 @ Tones
 XX42 TAI XAAU TONE MAI SAK
 x (tai viet tone mai ek - AABF)
 x (combining acute accent - 0301)
 XX43 TAI XAAU TONE MAI HOI
 x (tai viet tone mai tho - AAC1)
 x (combining hook above - 0309)
 @ Word ligature symbols
 XX44 TAI XAAU SYMBOL NUENG
 = one
 x (tai viet symbol nueng - AADC)
 XX45 TAI XAAU SYMBOL KUN
 = person
 x (tai viet symbol kon - AADB)
 XX46 TAI XAAU SYMBOL MA
 = horse
 XX47 TAI XAAU SYMBOL BEN
 = moon
 XX48 TAI XAAU SYMBOL MAN
 = the third person singular
 @ Punctuation
 XX49 TAI XAAU SYMBOL MAAI SUT
 XX4A TAI XAAU SYMBOL MAAI SAM
 x (tai viet symbol koi koi - AADF)
 XX4B TAI XAAU SYMBOL MAAI SAM KOAK
 x (tai viet symbol ho hoi - AADE)
 XX4C TAI XAAU SYMBOL MAAI SAM MAET
 XX4D TAI XAAU SYMBOL LAN KAN-1
 XX4E TAI XAAU SYMBOL LAN KAN-2
 XX4F TAI XAAU SYMBOL LAN KAN-3

The information in UnicodeData.txt is as follows.

XX00;TAI XAAU LETTER HIGH QO;Lo;0;L;;;;N;;;;;
 XX01;TAI XAAU LETTER LOW QO;Lo;0;L;;;;N;;;;;
 XX02;TAI XAAU LETTER HIGH KO;Lo;0;L;;;;N;;;;;
 XX03;TAI XAAU LETTER LOW KO;Lo;0;L;;;;N;;;;;

XX04;TAI XAAU LETTER HIGH XO;Lo;0;L;;;;N;;;;;
 XX05;TAI XAAU LETTER LOW XO;Lo;0;L;;;;N;;;;;
 XX06;TAI XAAU LETTER HIGH HO NGO;Lo;0;L;;;;N;;;;;
 XX07;TAI XAAU LETTER LOW NGO;Lo;0;L;;;;N;;;;;
 XX08;TAI XAAU LETTER HIGH TSO;Lo;0;L;;;;N;;;;;
 XX09;TAI XAAU LETTER LOW TSO;Lo;0;L;;;;N;;;;;
 XX0A;TAI XAAU LETTER HIGH SO;Lo;0;L;;;;N;;;;;
 XX0B;TAI XAAU LETTER LOW SO;Lo;0;L;;;;N;;;;;
 XX0C;TAI XAAU LETTER HIGH HO YO;Lo;0;L;;;;N;;;;;
 XX0D;TAI XAAU LETTER LOW YO;Lo;0;L;;;;N;;;;;
 XX0E;TAI XAAU LETTER HIGH TO;Lo;0;L;;;;N;;;;;
 XX0F;TAI XAAU LETTER LOW TO;Lo;0;L;;;;N;;;;;
 XX10;TAI XAAU LETTER HIGH THO;Lo;0;L;;;;N;;;;;
 XX11;TAI XAAU LETTER LOW THO;Lo;0;L;;;;N;;;;;
 XX12;TAI XAAU LETTER HIGH HO NO;Lo;0;L;;;;N;;;;;
 XX13;TAI XAAU LETTER LOW NO;Lo;0;L;;;;N;;;;;
 XX14;TAI XAAU LETTER HIGH PO;Lo;0;L;;;;N;;;;;
 XX15;TAI XAAU LETTER LOW PO;Lo;0;L;;;;N;;;;;
 XX16;TAI XAAU LETTER HIGH PHO;Lo;0;L;;;;N;;;;;
 XX17;TAI XAAU LETTER LOW PHO;Lo;0;L;;;;N;;;;;
 XX18;TAI XAAU LETTER HIGH HO MO;Lo;0;L;;;;N;;;;;
 XX19;TAI XAAU LETTER LOW MO;Lo;0;L;;;;N;;;;;
 XX1A;TAI XAAU LETTER HIGH FO;Lo;0;L;;;;N;;;;;
 XX1B;TAI XAAU LETTER LOW FO;Lo;0;L;;;;N;;;;;
 XX1C;TAI XAAU LETTER HIGH BO;Lo;0;L;;;;N;;;;;
 XX1D;TAI XAAU LETTER LOW BO;Lo;0;L;;;;N;;;;;
 XX1E;TAI XAAU LETTER HIGH HO LO;Lo;0;L;;;;N;;;;;
 XX1F;TAI XAAU LETTER LOW LO;Lo;0;L;;;;N;;;;;
 XX20;TAI XAAU LETTER HIGH HO;Lo;0;L;;;;N;;;;;
 XX21;TAI XAAU LETTER LOW HO;Lo;0;L;;;;N;;;;;
 XX22;TAI XAAU LETTER HIGH DO;Lo;0;L;;;;N;;;;;
 XX23;TAI XAAU LETTER LOW DO;Lo;0;L;;;;N;;;;;
 XX24;TAI XAAU LETTER HIGH HO VO;Lo;0;L;;;;N;;;;;
 XX25;TAI XAAU LETTER LOW VO;Lo;0;L;;;;N;;;;;
 XX26;TAI XAAU LETTER HIGH KHO;Lo;0;L;;;;N;;;;;
 XX27;TAI XAAU LETTER LOW KHO;Lo;0;L;;;;N;;;;;
 XX28;TAI XAAU LETTER HIGH TSHO;Lo;0;L;;;;N;;;;;
 XX29;TAI XAAU LETTER LOW TSHO;Lo;0;L;;;;N;;;;;
 XX2A;TAI XAAU LETTER HIGH HO NYO;Lo;0;L;;;;N;;;;;
 XX2B;TAI XAAU LETTER LOW NYO;Lo;0;L;;;;N;;;;;
 XX30;TAI XAAU MAAI KA-1;Lo;0;L;;;;N;;;;;
 XX31;TAI XAAU MAAI KA-2;Mn;220;NSM;;;;N;;;;;
 XX32;TAI XAAU MAAI KA-3;Mn;230;NSM;;;;N;;;;;
 XX33;TAI XAAU MAAI KA-4;Mn;230;NSM;;;;N;;;;;
 XX34;TAI XAAU MAAI KA-5;Mn;230;NSM;;;;N;;;;;
 XX35;TAI XAAU MAAI KI;Mn;230;NSM;;;;N;;;;;
 XX36;TAI XAAU MAAI KE;Mn;230;NSM;;;;N;;;;;
 XX37;TAI XAAU MAAI KAE;Lo;0;L;;;;N;;;;;
 XX38;TAI XAAU MAAI TSAM;Mn;230;NSM;;;;N;;;;;
 XX39;TAI XAAU MAAI KOA;Lo;0;L;;;;N;;;;;
 XX3A;TAI XAAU MAAI KU-1;Mn;220;NSM;;;;N;;;;;
 XX3B;TAI XAAU MAAI KU-2;Lo;0;L;;;;N;;;;;
 XX3C;TAI XAAU MAAI KU-3;Mn;230;NSM;;;;N;;;;;
 XX3D;TAI XAAU MAAI KUE;Mn;230;NSM;;;;N;;;;;
 XX3E;TAI XAAU MAAI KEI;Mn;230;NSM;;;;N;;;;;
 XX3F;TAI XAAU MAAI KAUE;Lo;0;L;;;;N;;;;;
 XX40;TAI XAAU MAAI KAU;Lo;0;L;;;;N;;;;;
 XX41;TAI XAAU MAAI KAI;Lo;0;L;;;;N;;;;;
 XX42;TAI XAAU TONE MAI SAK;Mn;230;NSM;;;;N;;;;;
 XX43;TAI XAAU TONE MAI HOI;Mn;230;NSM;;;;N;;;;;
 XX44;TAI XAAU SYMBOL NUENG;Lo;0;L;;;;N;;;;;
 XX45;TAI XAAU SYMBOL KUN;Lo;0;L;;;;N;;;;;
 XX46;TAI XAAU SYMBOL MA;Lo;0;L;;;;N;;;;;


```

XX47;TAI XAAU SYMBOL BEN;Lo;0;L;;;;N;;;;;
XX48;TAI XAAU SYMBOL MAN;Lo;0;L;;;;N;;;;;
XX49;TAI XAAU SYMBOL MAAI SUT;Po;0;L;;;;N;;;;;
XX4A;TAI XAAU DIGIT MAAI SAM;Po;0;L;;;;N;;;;;
XX4B;TAI XAAU DIGIT MAAI SAM KOAK;Po;0;L;;;;N;;;;;
XX4C;TAI XAAU DIGIT MAAI SAM MAET;Po;0;L;;;;N;;;;;
XX4D;TAI XAAU DIGIT LAN KAAAN-1;Po;0;L;;;;N;;;;;
XX4E;TAI XAAU DIGIT LAN KAAAN-2;Po;0;L;;;;N;;;;;
XX4F;TAI XAAU DIGIT LAN KAAAN-3;Po;0;L;;;;N;;;;;

```

The Tai XaaU script is a newly proposed script for ISO 15924. The suggested code is `Taxa`.

Script=Taxa

The information in `IndicPositionalCategory.txt` is as follows.

```

# Indic_Positional_Category=Right
XX30      ; Right      # Lo      TAI XAAU MAAI KA-1
XX39      ; Right      # Lo      TAI XAAU MAAI KOA
XX40      ; Right      # Lo      TAI XAAU MAAI KAU

# Indic_Positional_Category=Visual_Order_Left
XX37      ; Visual_Order_Left  # Lo      TAI XAAU MAAI KAE
XX3A      ; Visual_Order_Left  # Lo      TAI XAAU MAAI KU-2
XX3E..XX40 ; Visual_Order_Left  # Lo [3] TAI XAAU MAAI KAUE..TAI XAAU MAAI KAI

# Indic_Positional_Category=Top
XX32..XX36 ; Top          # Mn [5] TAI XAAU MAAI KA-3..TAI XAAU MAAI KE
XX38      ; Top          # Mn      TAI XAAU MAAI TSAM
XX3C..XX3E ; Top          # Mn [3] TAI XAAU MAAI KU-3..TAI XAAU MAAI KEI
XX42..XX43 ; Top          # Mn [2] TAI XAAU TONE MAI SAK..TAI XAAU TONE
      MAI HOI

# Indic_Positional_Category=Bottom
XX31      ; Bottom      # Mn      TAI XAAU MAAI KA-2
XX3A      ; Bottom      # Mn      TAI XAAU MAAI KU-1

```

The information in `IndicSyllableCategory.txt` is as follows.

```

# Indic_Syllabic_Category=Vowel_Dependent
XX30      ; Vowel_Dependent # Lo      TAI XAAU MAAI KA-1
XX31      ; Vowel_Dependent # Mn      TAI XAAU MAAI KA-2
XX32..XX36 ; Vowel_Dependent # Mn [5] TAI XAAU MAAI KA-3..TAI XAAU MAAI KE
XX37      ; Vowel_Dependent # Lo      TAI XAAU MAAI KAE
XX38      ; Vowel_Dependent # Mn      TAI XAAU MAAI TSAM
XX39      ; Vowel_Dependent # Lo      TAI XAAU MAAI KOA
XX3A      ; Vowel_Dependent # Mn      TAI XAAU MAAI KU-1
XX3B      ; Vowel_Dependent # Lo      TAI XAAU MAAI KU-2
XX3C..XX3E ; Vowel_Dependent # Mn [3] TAI XAAU MAAI KU-3..TAI XAAU MAAI KEI
XX3F..XX41 ; Vowel_Dependent # Lo [3] TAI XAAU MAAI KAUE..TAI XAAU MAAI KAI

# Indic_Syllabic_Category=Consonant
XX00..XX2B ; Consonant      # Lo [44] TAI XAAU LETTER HIGH QO..TAI XAAU LETTER
LOW NYO

# Indic_Syllabic_Category=Tone_Mark
XX42..XX43 ; Tone_Mark      # Mn [2] TAI XAAU TONE MAI SAK..TAI XAAU TONE
      MAI HOI

```

The information in `LineBreak.txt` is as follows.

```

XX00..XX2B ; SA      # Lo [44] TAI XAAU LETTER HIGH QO..TAI XAAU LETTER LOW NYO
XX30      ; SA      # Lo      TAI XAAU MAAI KA-1
XX31      ; SA      # Mn      TAI XAAU MAAI KA-2

```

```

XX32..XX36 ; SA # Mn [5] TAI XAAU MAAI KA-3..TAI XAAU MAAI KE
XX37 ; SA # Lo TAI XAAU MAAI KAE
XX38 ; SA # Mn TAI XAAU MAAI TSAM
XX39 ; SA # Lo TAI XAAU MAAI KOA
XX3A ; SA # Mn TAI XAAU MAAI KU-1
XX3B ; SA # Lo TAI XAAU MAAI KU-2
XX3C..XX3E ; SA # Mn [3] TAI XAAU MAAI KU-3..TAI XAAU MAAI KEI
XX3F..XX41 ; SA # Lo [3] TAI XAAU MAAI KAUE..TAI XAAU MAAI KAI
XX42..XX43 ; SA # Mn [2] TAI XAAU TONE MAI SAK..TAI XAAU TONE MAI HOI
XX44..XX48 ; SA # Lo [5] TAI XAAU SYMBOL NUENG..TAI XAAU SYMBOL MAN
XX49..XX4F ; SA # Po [7] TAI XAAU SYMBOL MAAI SUT..TAI XAAU SYMBOL LAN KAN-3

```

The information in PropList.txt is as follows.

```

XX49..XX4F ; Terminal_Punctuation # Po [7] TAI XAAU SYMBOL MAAI SUT..TAI XAAU
SYMBOL LAN KAN-3
XX49..XX4F ; Sentence_Terminal # Po [7] TAI XAAU SYMBOL MAAI SUT..TAI XAAU
SYMBOL LAN KAN-3

XX31 ; Other_Alphabetic # Mn TAI XAAU MAAI KA-2
XX32..XX36 ; Other_Alphabetic # Mn [5] TAI XAAU MAAI KA-3..TAI XAAU MAAI KE
XX38 ; Other_Alphabetic # Mn TAI XAAU MAAI TSAM
XX3A ; Other_Alphabetic # Mn TAI XAAU MAAI KU-1
XX3C..XX3E ; Other_Alphabetic # Mn [3] TAI XAAU MAAI KU-3..TAI XAAU MAAI KEI

XX42..XX43 ; Diacritic # Mn [2] TAI XAAU TONE MAI SAK..TAI XAAU TONE
MAI HOI

XX37 ; Logical_Order_Exception # Lo TAI XAAU MAAI KAE
XX3A ; Logical_Order_Exception # Lo TAI XAAU MAAI KU-2
XX3E..XX40 ; Logical_Order_Exception # Lo [3] TAI XAAU MAAI KAUE..TAI
XAAU MAAI KAI

```

The information in VerticalOrientation.txt is as follows.

```

XX00..XX2B ; R # Lo [44] TAI XAAU LETTER HIGH QO..TAI XAAU LETTER LOW NYO
XX30 ; R # Lo TAI XAAU MAAI KA-1
XX31 ; R # Mn TAI XAAU MAAI KA-2
XX32..XX36 ; R # Mn [5] TAI XAAU MAAI KA-3..TAI XAAU MAAI KE
XX37 ; R # Lo TAI XAAU MAAI KAE
XX38 ; R # Mn TAI XAAU MAAI TSAM
XX39 ; R # Lo TAI XAAU MAAI KOA
XX3A ; R # Mn TAI XAAU MAAI KU-1
XX3B ; R # Lo TAI XAAU MAAI KU-2
XX3C..XX3E ; R # Mn [3] TAI XAAU MAAI KU-3..TAI XAAU MAAI KEI
XX3F..XX41 ; R # Lo [3] TAI XAAU MAAI KAUE..TAI XAAU MAAI KAI
XX42..XX43 ; R # Mn [2] TAI XAAU TONE MAI SAK..TAI XAAU TONE MAI HOI
XX44..XX48 ; R # Lo [5] TAI XAAU SYMBOL NUENG..TAI XAAU SYMBOL MAN
XX49..XX4F ; R # Po [7] TAI XAAU SYMBOL MAAI SUT..TAI XAAU SYMBOL LAN KAN-3

```

The information in DoNotEmit.txt is as follows.

```

# Tai_Xaau_Word_Ligature:
# Sequences that look like a Tai Xaau word ligature but should be avoided
# in representing that word ligature.

XX19 XX32; XX48; Tai_Xaau_Word_Ligature # TAI XAAU LETTER LOW MO, TAI XAAU MAAI KA-3;
TAI XAAU SYMBOL MAN

```

The information in confusables.txt (UTR #39) is as follows.

```

XX37 ; XX40 XX40 ; MA #* ( “ → ” ) TAI XAAU MAAI KAE → TAI XAAU MAAI KAU, TAI XAAU
MAAI KAU #

```

Reference

Compilation Office of Honghe Hani and Yi Autonomous Prefecture 红河哈尼族彝族自治州民族志编写办公室, (1989). *Ethnography of Honghe Hani and Yi Autonomous Prefecture in Yunnan Province* / 云南省红河哈尼族彝族自治州民族志. Kunming / 昆明: Yunnan University Press / 云南大学出版社.

DAO Jie 刀洁, (2016). “Phonological Difference between Jinping Tai Script and Jinping Tai Language” / 金平傣文和金平傣语的音系差异, in: *Protection and Inheritance of Language Resources* / 语言资源的保护与传承. Beijing / 北京: Minzu Publishing House / 民族出版社.

DAO Jie 刀洁, ZHOU Yan 周焱, WANG Yuqing 王玉清, (2019). “The standardization of the phonetics of Jinping Dai writing system in Yunnan” / 云南金平傣文语音的规范化问题, in: *Language Plannology Studies* / 语言规划学研究, No. 2: pp. 57–68.

HE Bingkun 何炳坤, (1999). “Jinping Dai Script” / 金平傣文, in: HE Lifeng 和丽峰, XIONG Yuyou 熊玉有. *Outline of Yunnan Minzu Scripts* / 云南少数民族文字概要. Kunming / 昆明: Yunnan Minzu Publishing House / 云南民族出版社.

HE Shaoyin 和少英, DAO Jie 刀洁, (2007). *The Catchers in their Homeland on the Boundary Line — the Society and Culture of Jinping Dai Minzu* / 守望国境线上的家园——金平傣族的社会文化. Kunming / 昆明: Yunnan University Publishing House / 云南大学出版社.

LI Mengdie 李梦蝶, (2020). *A Comparative Study on the Phonetics of Thai and Jinping Dai Languages* / 泰语和金平傣语的语音比较研究. Kunming / 昆明: Yunnan Minzu University / 云南民族大学, Master Thesis.

LUO Wenbin 罗文斌 (chief editor), (2017). *金平傣文教程*. Honghe / 红河: Cultural Center of Jinping Miao, Yao and Dai Autonomous County / 金平苗族瑶族傣族自治县文化馆.

Pittayawat PITTAYAPORN, (2009). *The Phonology of Proto-Tai*. New York: Cornell University, PhD Dissertation.

Planning Office for the Collation and Publication of Minzu Ancient Literature in Yunnan Province 云南省少数民族古籍整理出版规划办公室, (2020). *Collection of Yunnan Minzu Rare Ancient Literature* / 云南少数民族古籍珍本集成. Kunming / 昆明: Yunnan People's Publishing House / 云南人民出版社, Vol. 86.

<http://www.kaom.net/min_z_kamtai.php>

<<https://zhongguoyuyan.cn/province/62/侗台语族>>

Appendix A Glyphs comparison

A.1 Glyphs in ancient literatures

Below is a list of the glyphs used in Tai Xaa script, where “—” means unattested. Glyphs in the second to fourth columns are from [Dao et al. \(2019\)](#) [Dao], [Luo \(2017\)](#) [Luo] and [COHAP \(1989\)](#) [COHAP]. And glyphs in the remaining five columns are from ancient literatures in [POCP \(2020\)](#), they are:

- **DD:** *Dōngdēng Gùshì* 东登故事, “Story of Dongdeng (person name)”;
- **HL:** *Hūnlúo Yǔ Nánwō* 婚罗与喃窝, “Hunluo (person name) and Nanwo (person name)”;
- **ZL:** *Zhúlóu Shíwù* 竹楼实务, “Practice of Building Bamboo Houses”;
- **AXL:** *Àixiànglún* 艾相伦, “Aixianglun (person name)”;
- **Others:** *Cǎiyào Pǔjì* 采药谱记, “Records of Herb Collection” and *Qínggē* 情歌, “Love Song”.

Glyph	Dao	Luo	COHAP	DD	HL	ZL	AXL	Others
9	9	9	9					
9	9	9	9					—
n	n	n	n					
n	n	n	n					
16	16	16	16					
7	7	7	7					
916	916	916	916					
6	6	6	6					
5	5	5	5					
m	m	m	m					
m	m	m	m					
m	m	m	m					
96	96	96	96					

				—	—	—		—
			—	—	—	—		—
			—					
			—					
			—			—		—
			—	—	—	—		—
	—	—	—	—	—	—		—
			—	—	—	—		—
			—	—	—	—		—
			—					—
			—					—
			—					
	—	—	—		—	—		—
			—	—	—	—		—
			—	—	—	—		—

A.2 Glyphs in modern literatures

Below is a list of the glyphs used in Tai Xaa script. Glyphs in the second to sixth columns are from:

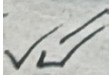
- **DD:** *Dōngdēng* 冬灯, “Dongdeng (person name)”;
- **CS:** Capitaine SILVESTRE, (1918), “*Les Thai Blancs de Phong-Tho*”, in: *Bulletin de l’École Française d’Extrême-Orient*, Vol. 18, No. 4: pp. 1–56.
- **GM:** Georges MINOT, (1940), “*Dictionnaire Tay Blanc Français*”, in: *Bulletin de l’École française d’Extrême-Orient*, Vol. 40, No. 1: pp. 1–237.

- **MF:** Michel FERLUS, (2006), “[Les écritures thai du Vietnam](#)”, in: *Cahiers de Linguistique – Asie Orientale*, Vol. 35, No. 2: pp. 209–239.
- **HDL:** HE Bingkun 何炳坤, DAO Jiafan 刀家凡, LUO Jun 罗军, (1996), *Jinping Dai Textbook (for Literacy)* / 金平傣文课本 (扫盲用). Honghe / 红河: Education Management Committee of Mengla Town, Jinping Miao, Yao and Dai Autonomous County / 金平苗族瑶族傣族自治县勐拉乡教育管理委员会.

Glyph	DD	CS	GM	MF	HDL
๑					
๒					
๓					
๔					
๕					
๖					
๗					
๘					
๙					
๐					
๑๐					
๑๑					
๑๒					
๑๓					
๑๔					
๑๕					
๑๖					
๑๗					
๑๘					
๑๙					
๒๐					

91					
1					
2					
3					
4					
5					
6					
7					
8					
9					
0					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					

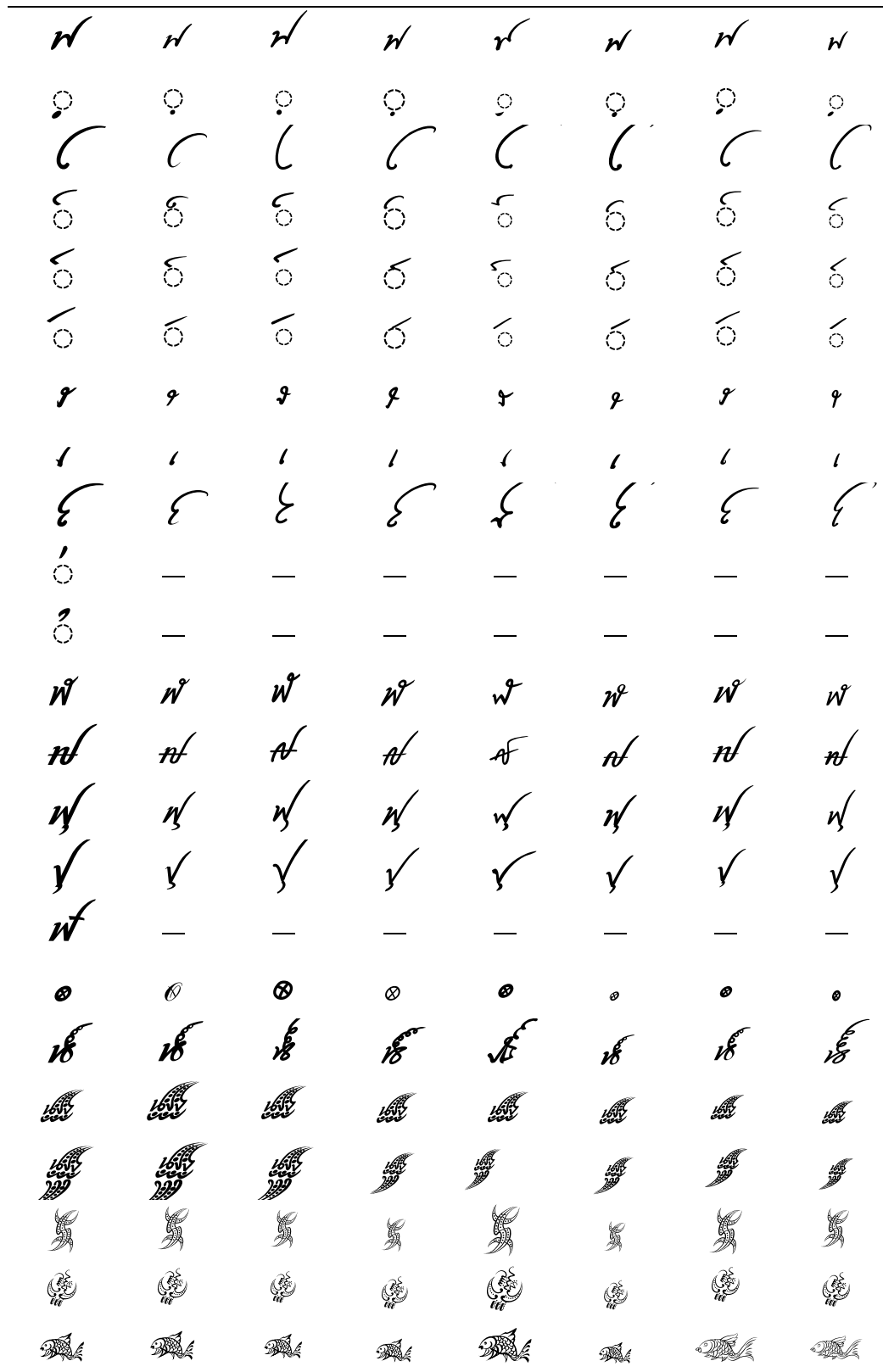
4	4	4	4	4	4
913	913	913	913	913	913
113	113	113	113	113	113
,	,	,	,	,	,
°	°	°	°	°	°
17	17	17	17	17	17
16	16	16	16	16	16
✓	✓	✓	✓	✓	✓
○	○	○	○	○	○
○	○	○	○	○	○
○	○	○	○	○	○
○	○	○	○	○	○
○	○	○	○	○	○
11	11	11	11	11	11
°	°	°	°	°	°
W	W	W	W	W	W
°	°	°	°	°	°
○	○	○	○	○	○
○	○	○	○	○	○
○	○	○	○	○	○
○	○	○	○	○	○
8	8	8	8	8	8
1	1	1	1	1	1
3	3	3	3	3	3
○	○	○	○	○	○

?	—	—	—	—	—
ŵ		—	ŵ	—	
nt		—	nt	—	
ŵ		—	—	—	
ŵ		—	—	—	
nt		—	—	—	—
o		—	—	—	
ŵ		—	ŵ	—	—
		—		—	—
		—	—	—	—
		—	—	—	—
		—	—	—	—
		—	—	—	—

A.3 Glyphs in computerize fonts

Below is a list of the glyphs in several computerize fonts used in Tai Xaau script. The names of fonts are: Xiaorui Regular (小睿体, made in 2018), Huangtao (黄韬体, made in 2019, HT), Huanghe (黄禾体, made in 2022, HH), Xiaosong (小宋体, made in 2022, XS), Huataiwen (华泰温体, made in 2022, HTW), Yinzong (尹宗体, made in 2023, YZ), Xiaorui Light (小睿体细体, made in 2024), Bobo (波波体, made in 2024, BB).

Glyph Xiaorui	HT	HH	XS	HTW	YZ	Xiaorui Light	BB
o	o	o	o	o	o	o	o
o	o	o	o	o	o	o	o
n	n	n	n	n	n	n	n
nt	nt	nt	nt	nt	nt	nt	nt



	xx0	xx1	xx2	xx3	xx4
0	 xx00	 xx10	 xx20	 xx30	 xx40
1	 xx01	 xx11	 xx21	 xx31	 xx41
2	 xx02	 xx12	 xx22	 xx32	 xx42
3	 xx03	 xx13	 xx23	 xx33	 xx43
4	 xx04	 xx14	 xx24	 xx34	 xx44
5	 xx05	 xx15	 xx25	 xx35	 xx45
6	 xx06	 xx16	 xx26	 xx36	 xx46
7	 xx07	 xx17	 xx27	 xx37	 xx47
8	 xx08	 xx18	 xx28	 xx38	 xx48
9	 xx09	 xx19	 xx29	 xx39	 xx49
A	 xx0A	 xx1A	 xx2A	 xx3A	 xx4A
B	 xx0B	 xx1B	 xx2B	 xx3B	 xx4B
C	 xx0C	 xx1C		 xx3C	 xx4C
D	 xx0D	 xx1D		 xx3D	 xx4D
E	 xx0E	 xx1E		 xx3E	 xx4E
F	 xx0F	 xx1F		 xx3F	 xx4F

Consonants

xx00	ᵛ	TAI XAAU LETTER HIGH QO → AAAE ᵛ tai viet letter low o
xx01	ᵛ	TAI XAAU LETTER LOW QO → AAAF ᵛ tai viet letter high o
xx02	ᵛ	TAI XAAU LETTER HIGH KO → AA80 ᵛ tai viet letter low ko
xx03	ᵛ	TAI XAAU LETTER LOW KO
xx04	ᵛ	TAI XAAU LETTER HIGH XO
xx05	ᵛ	TAI XAAU LETTER LOW XO
xx06	ᵛ	TAI XAAU LETTER HIGH HO NGO
xx07	ᵛ	TAI XAAU LETTER LOW NGO → AA89 ᵛ tai viet letter high ngo
xx08	ᵛ	TAI XAAU LETTER HIGH TSO → AA8A ᵛ tai viet letter low co
xx09	ᵛ	TAI XAAU LETTER LOW TSO → AA8B ᵛ tai viet letter high co
xx0A	ᵛ	TAI XAAU LETTER HIGH SO
xx0B	ᵛ	TAI XAAU LETTER LOW SO
xx0C	ᵛ	TAI XAAU LETTER HIGH HO YO
xx0D	ᵛ	TAI XAAU LETTER LOW YO → AAA5 ᵛ tai viet letter high yo
xx0E	ᵛ	TAI XAAU LETTER HIGH TO → AA94 ᵛ tai viet letter low to
xx0F	ᵛ	TAI XAAU LETTER LOW TO → AA95 ᵛ tai viet letter high to
xx10	ᵛ	TAI XAAU LETTER HIGH THO → AA96 ᵛ tai viet letter low tho
xx11	ᵛ	TAI XAAU LETTER LOW THO
xx12	ᵛ	TAI XAAU LETTER HIGH HO NO → AA98 ᵛ tai viet letter low no
xx13	ᵛ	TAI XAAU LETTER LOW NO → AA99 ᵛ tai viet letter high no
xx14	ᵛ	TAI XAAU LETTER HIGH PO
xx15	ᵛ	TAI XAAU LETTER LOW PO
xx16	ᵛ	TAI XAAU LETTER HIGH PHO
xx17	ᵛ	TAI XAAU LETTER LOW PHO
xx18	ᵛ	TAI XAAU LETTER HIGH HO MO
xx19	ᵛ	TAI XAAU LETTER LOW MO
xx1A	ᵛ	TAI XAAU LETTER HIGH FO
xx1B	ᵛ	TAI XAAU LETTER LOW FO
xx1C	ᵛ	TAI XAAU LETTER HIGH BO → AA9A ᵛ tai viet letter low bo
xx1D	ᵛ	TAI XAAU LETTER LOW BO → AA9B ᵛ tai viet letter high bo
xx1E	ᵛ	TAI XAAU LETTER HIGH HO LO → AAA8 ᵛ tai viet letter low lo
xx1F	ᵛ	TAI XAAU LETTER LOW LO
xx20	ᵛ	TAI XAAU LETTER HIGH HO → AAAC ᵛ tai viet letter low ho
xx21	ᵛ	TAI XAAU LETTER LOW HO → AAAD ᵛ tai viet letter high ho
xx22	ᵛ	TAI XAAU LETTER HIGH DO
xx23	ᵛ	TAI XAAU LETTER LOW DO → AA93 ᵛ tai viet letter high do
xx24	ᵛ	TAI XAAU LETTER HIGH HO VO → AAAA ᵛ tai viet letter low vo
xx25	ᵛ	TAI XAAU LETTER LOW VO → AAAB ᵛ tai viet letter high vo
xx26	ᵛ	TAI XAAU LETTER HIGH KHO
xx27	ᵛ	TAI XAAU LETTER LOW KHO → AA83 ᵛ tai viet letter high kho
xx28	ᵛ	TAI XAAU LETTER HIGH TSHO → AA8C ᵛ tai viet letter low cho
xx29	ᵛ	TAI XAAU LETTER LOW TSHO → AA8D ᵛ tai viet letter high cho
xx2A	ᵛ	TAI XAAU LETTER HIGH HO NYO
xx2B	ᵛ	TAI XAAU LETTER LOW NYO

Vowels

xx30	ᵛ	TAI XAAU MAAI KA-1 → AAB1 1 tai viet vowel aa
xx31	ᵛ	TAI XAAU MAAI KA-2
xx32	ᵛ	TAI XAAU MAAI KA-3
xx33	ᵛ	TAI XAAU MAAI KA-4 → AAB0 ᵛ tai viet mai kang
xx34	ᵛ	TAI XAAU MAAI KA-5 → AABE ᵛ tai viet vowel am
xx35	ᵛ	TAI XAAU MAAI KI → AAB2 ᵛ tai viet vowel i
xx36	ᵛ	TAI XAAU MAAI KE → AAB8 ᵛ tai viet vowel ia
xx37	ᵛ	TAI XAAU MAAI KAE • precedes consonant in visual order → AAB5 ᵛ tai viet vowel e
xx38	ᵛ	TAI XAAU MAAI TSAM
xx39	ᵛ	TAI XAAU MAAI KOA → AABA ᵛ tai viet vowel ua
xx3A	ᵛ	TAI XAAU MAAI KU-1 → AAB4 ᵛ tai viet vowel u
xx3B	ᵛ	TAI XAAU MAAI KU-2 • precedes consonant in visual order → AAB6 ᵛ tai viet vowel o
xx3C	ᵛ	TAI XAAU MAAI KU-3
xx3D	ᵛ	TAI XAAU MAAI KUE → AAB3 ᵛ tai viet vowel ue
xx3E	ᵛ	TAI XAAU MAAI KEI → AAB7 ᵛ tai viet mai khít
xx3F	ᵛ	TAI XAAU MAAI KAUE • precedes consonant in visual order → AABB ᵛ tai viet vowel aue
xx40	ᵛ	TAI XAAU MAAI KAU • precedes and follows consonant in visual order → AAB9 ᵛ tai viet vowel uea
xx41	ᵛ	TAI XAAU MAAI KAI • precedes consonant in visual order → AABC ᵛ tai viet vowel ay

Tones

xx42	ᵛ	TAI XAAU TONE MAI SAK → AABF ᵛ tai viet tone mai ek → 0301 ᵛ combining acute accent
xx43	ᵛ	TAI XAAU TONE MAI HOI → AAC1 ᵛ tai viet tone mai tho → 0309 ᵛ combining hook above

Word ligature symbols

xx44	ᵛ	TAI XAAU SYMBOL NUENG = one → AADC ᵛ tai viet symbol nueng
xx45	ᵛ	TAI XAAU SYMBOL KUN = person → AADB ᵛ tai viet symbol kon
xx46	ᵛ	TAI XAAU SYMBOL MA = horse
xx47	ᵛ	TAI XAAU SYMBOL BEN = moon
xx48	ᵛ	TAI XAAU SYMBOL MAN = the third person singular

Punctuation

xx49	ᵛ	TAI XAAU SYMBOL MAAI SUT
xx4A	ᵛ	TAI XAAU SYMBOL MAAI SAM → AADF ᵛ tai viet symbol koi koi
xx4B	ᵛ	TAI XAAU SYMBOL MAAI SAM KOAK → AADE ᵛ tai viet symbol ho hoi
xx4C	ᵛ	TAI XAAU SYMBOL MAAI SAM MAET
xx4D	ᵛ	TAI XAAU SYMBOL LAN KAAAN-1
xx4E	ᵛ	TAI XAAU SYMBOL LAN KAAAN-2
xx4F	ᵛ	TAI XAAU SYMBOL LAN KAAAN-3

Figures

ㄴㅇ 44 辅音字母表

第二课 复合元音

一、韵尾辅音字母 (8 个)

y	'yo	'vo	'go	'no	'mo	ka	rr	po
y	'v	ŋ	n	m	k	rr		p

在学习复合元音之前请大家熟记以上八个辅音字母。
八个作为韵尾使用的辅音字母中前四个为高音质后、四个为低音质。

二、单元音 (9 个)

/a/	/i/	/e/	/ɛ/	/u/	/ʌ/	/ɔ/	/o/	/u/
-----	-----	-----	-----	-----	-----	-----	-----	-----

三、拼读规则

单元音

mai ka
a

→

韵尾辅音

me 'yo
'kay 'yo

→

ay
mai ay

复合元音字母表

y	'yo	'vo	ŋ	n	m	k	rr	p
mai ka	'kay 'yo	'kaw 'vo	kup 'go	'kan 'no	'kam 'mo	gak go	gat rro	gap po
a	ay	aw	aj	an	am	ak	at	ap
mai a	ay 'yo	aw 'vo	aj 'go	'an 'no	'am 'mo	ak ka	ot rro	op po
o	oy	ow	oj	on	om	ok	ot	op
mai o	oy 'yo	ow 'vo	oj 'go	'on 'no	'om 'mo	ok ka	ot rro	op po
u	uy	uw	uj	un	um	uk	ut	up
mai u	uy 'yo	uw 'vo	uj 'go	'un 'no	'um 'mo	uk ka	ut rro	up po

Figure 1 Tai Xaaü textbook, from top to bottom:
(a) Consonant letters and vowel letters; (b) a list of finals
Source: [Luo \(2017\)](#), pages 1–2, 7–8

表 8 规范声母

高音组			低音组		
p			p		
n	h	sh	n	h	h
h	h	sh	h	h	h
m	h	sh	m	h	h
h	h	sh	h	h	h
h	sh	h	h	h	h
h	h	h	h	h	h
h	h	h	h	h	h

4. 金平傣文规范韵母

以 83 个韵母作为金平傣文的规范韵母。包含 9 个单元音、14 个双元音、60 个带辅音韵尾的韵母（简称“辅音尾韵母”）。列表 9。

表 9 规范韵母

单元音		o	o	o	u	u	u	u	u	u
双元音	uo									
	uo	uo				uo	uo	uo	uo	uo
	uo	uo	uo	uo	uo	uo	uo	uo	uo	uo
辅音尾韵母	uo	uo	uo	uo	uo	uo	uo	uo	uo	uo
	uo	uo	uo	uo	uo	uo	uo	uo	uo	uo
	uo	uo	uo	uo	uo	uo	uo	uo	uo	uo
	uo	uo	uo	uo	uo	uo	uo	uo	uo	uo
	uo	uo	uo	uo	uo	uo	uo	uo	uo	uo
	uo	uo	uo	uo	uo	uo	uo	uo	uo	uo

5. 金平傣文规范声调

以六个调位作为金平傣文的 6 个声调，3 个声调符号作为规范调符。列表 10。

表 10 规范声调

调位	一/33/	二/55/	三/25/	四/43/	五/52/	六/31/
调值	33	55	25	43	52	31
调符						
例词	h	h	h	h	h	h

Figure 3 Tai Xaau orthography:
 Initials [table 8], finals [table 9] and tones [table 10]
 Source: [Dao et al. \(2019\)](#), page 67

/55/、/25/、/43/、/52/、/31/。其中调位/55/包含了舒声调第2调的调值[55]和入声调第7调的调值[55]；调位/43/包含了舒声调第4调的调值[43]和入声调第8调的调值[43]。

四、特殊符号

金平傣文有一些符号较为特殊，既不是声母，也不是韵母，而是一种固定的特殊符号，专门用来书写某个词或表示句末停顿的符号。这样的符号不多，可分两类：一种是固定字符；另一种是句末符号。

(一) 固定字符

固定字符是指用一个特定的符号来表示某一个词的固定书写形式，构成独立使用的独体字。这种固定字符不多，传统文字中只出现在“一、人、月亮、马”4个词中。即

𐄓 [nɯŋ⁵²] 一 𐄔 [kun⁴³] 人
𐄕 [bən³⁸] 月亮 𐄖 [ma³¹] 马

计算机字体根据需要增加了2个类似的独体字。即

𐄗 [lu⁴³] 孩子 𐄘 [la²⁵] 迟

上述的独体字也可以用声韵调拼合的方式书写，如：𐄓 (一)、𐄔 (人)、𐄕 (月亮)、𐄖 (马)、𐄗 (孩子)、𐄘 (迟)。使用时可根据需要选择使用独特字或声韵调拼合的字。

(二) 句末符号

金平傣文一般不使用标点符号，每一句话结束时就在后面标注一个句末符号。但并没有固定用哪一个字符，大多因人而异。有的字符与其说是字，不如说是图画，有的字形类似象形文字。句末符号有如下6个。即



其中，前5个均统称为[ma: i³³ sam³¹]，有“结束符号”之意，通常用在长篇的故事或诗歌里。最后1个称为[ma: i³³ mət⁵⁵]，有“完成符号”之意，用在句子之后，表示一句话的完成。

五、结语

综上所述，金平傣文是一种源于印度梵文字母系统的拼音文字，也是跨境自称“傣端”或“傣皓”的傣族白傣支系使用的民族文字。该文字由于创制的年代久远，与现行口语在声母、韵母、声调等方面都出现了一定的差异。加之长期以来均通过手抄方式流传于民间，字体形式因人而异，既没有规范的字体，也没有计算机字体，这对跨境傣族语言文化的传承和发展带来了不利的影响。文字是一个民族进步的标志性特征之一，它对推动民族文化的发展起到积极的作用。尤其是当今数字化、信息化的时代，文字所发挥的作用更加具有不可替代性。为了能更好保护与传承优秀的民族文化，本文通过对金平傣文的实地调查，并根据现行口语的语音特征，对其语音的规范化问题进行了探讨，并提出金平傣文语音规范化的基本原则和标准。具体归纳总结表8。

1. 金平傣文规范字体

以金平县融媒体中心罗文斌同志所设计的金平傣文计算机字体小睿体作为金平傣文数字化、信息化的规范字体。该字体包含了在传统文字符号基础上所设计的字体和为完整记录现行语言而新增的文字符号字体。

2. 金平傣文标准音

以金平县勐拉镇旧勐村的白傣语语音作为金平傣文的标准音。该村落的历史文化、地理位置、语音特点，符合标准选择的一般原则。

3. 金平傣文规范声母

以22个辅音音位44个傣文字母作为金平傣文的规范声母。列表8。

Figure 4 Tai Xaau orthography:
Symbols [section 4.1], punctuation [section 4.2]
Source: Dao et al. (2019), page 68

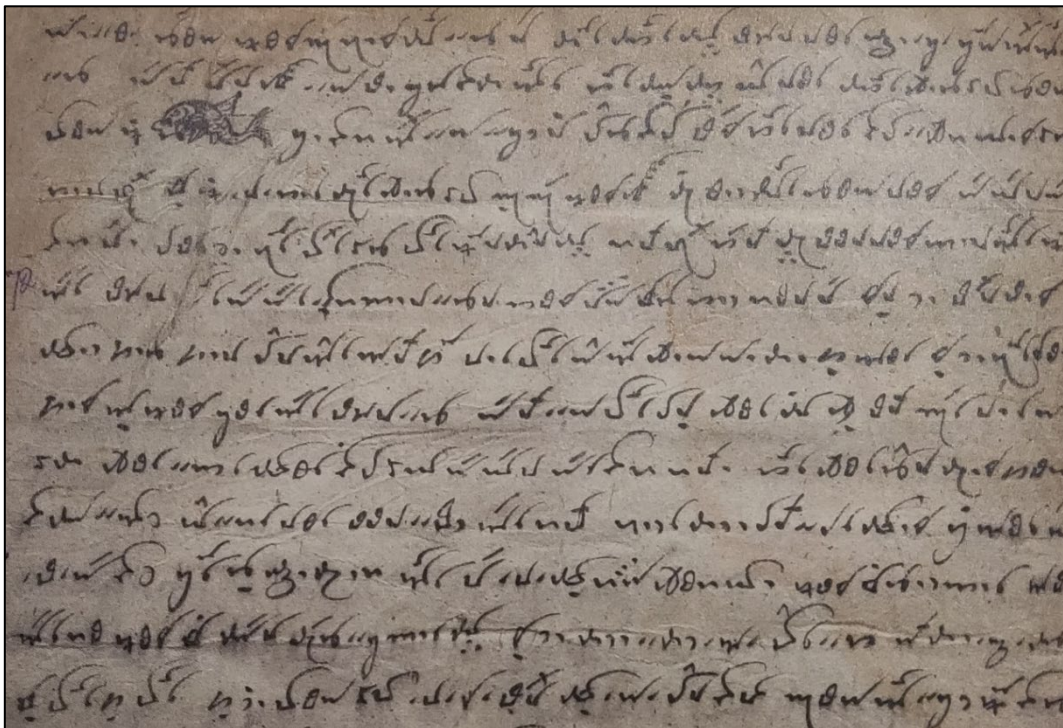
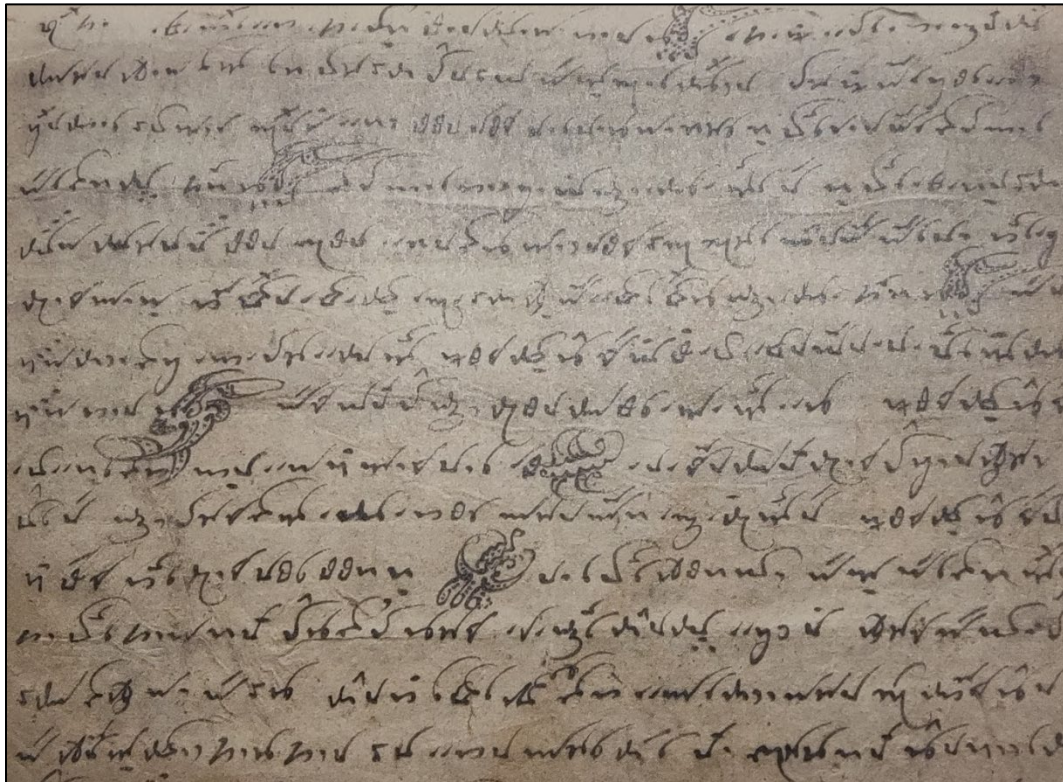


Figure 5 Tai Xaau literature, from top to bottom:
 (a) Punctuation ,  and ; (b) Punctuation 
 Source: “Aixianglun”, in: [POCP \(2020\)](#), pages 329–330

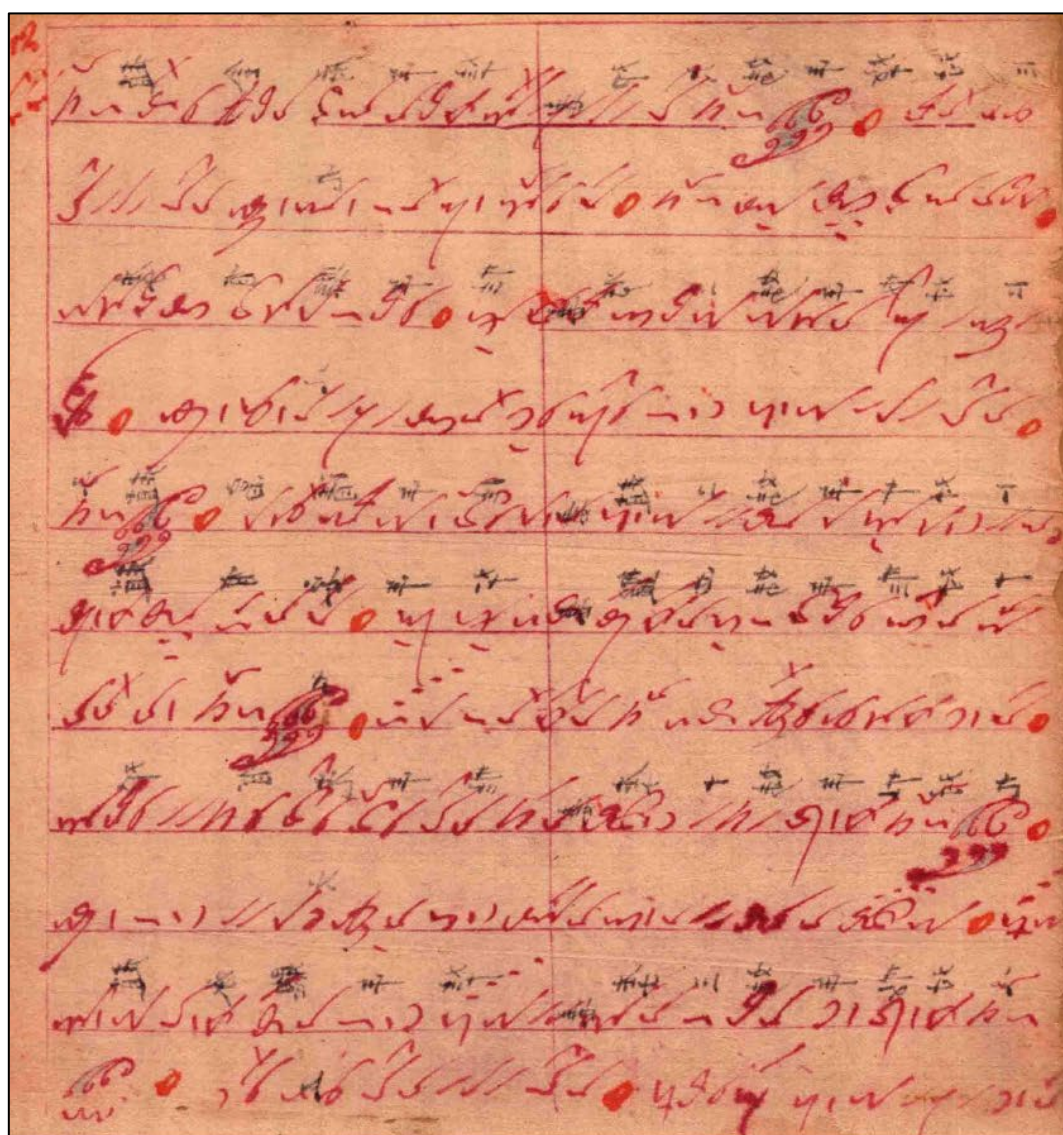


Figure 6 Tai Xaa literature: Punctuation  and 
Source: “Hunluo and Nanwo”, in: [POCP \(2020\)](#), page 61



Figure 7 Teaching Tai Xaau script, from top to bottom:
 (a) Teaching the Tai Xaau script, note that the blackboard shows six consonant letters with “ho” in their character names; (b) Students copy the letters from the textbook
 Source: Provided by LUO Wenbin



Figure 8 Tai Xaau script on slogans, from top to bottom:
 (a) Slogan in the school; (b) Slogan in the garden; (c) Slogan on the wall
 Source: Provided by LUO Wenbin

内部资料 注意保管 云南省少数民族语文指导工作委员会办公室 会议纪要 第 12 期 云南省民语委办公室 2024 年 9 月 18 日 金平傣文字符编码国际标准研制申报研讨会纪要 2024年9月10日，云南省民语委办公室和云南省红河州金平县民族宗教局在云南省红河州金平县组织召开金平傣文字符编码国际标准研制申报研讨会，中国电子技术标准化研究院、云南省民语委办公室、云南民族大学、金平县民族宗教局、金平县融媒体中心、金平县傣学研究会等单位的16位专家参加会议。会议听取了金平傣文字符编码国际标准方案介绍，并就相关问题进行了讨论，现纪要如下。 一、关于研制申报的重要性和必要性 会议认为，金平傣文作为四种傣族文字之一，字形符号独特，使用范围涵盖中国云南红河州金平县及相邻地区，越南西北部及老挝部分地区，使用人口有50万左右。历史上形成了一定数量的	文献，至今仍有不少人学习使用。目前，德宏傣文、西双版纳傣文、越南傣越文等已研制申报了国际标准，但代替不了金平傣文。研制申报金平傣文字符编码国际标准，有利于实现金平傣文的信息化，有利于抢救保护珍贵的傣族优秀传统文化，有利于维护国家文化安全和主权，具有重要现实意义，应当抓紧进行，积极推进。 二、关于文字名称 金平傣文也称傣端文、傣暗文、傣密文、白傣文等。研制申报字符编码国际标准，需要统一名称。大家认为，“金平傣文”是国内通俗称呼，存在一定的局限性。为充分尊重本民族意愿，会议决定采用“傣暗文”这一名称，可涵盖中国金平地区以及越南等自称“傣暗”的傣族支系，范围广，认可度高，易于接受。正式方案名称为：“傣暗文字符国际标准编码方案”。 三、关于字符 主要涉及以下几个方面： （一）标点符号 传统金平傣文用图画表示句子或段落结束，如“鱼”形等，形象生动，但笔画复杂，不便书写，方案中收录了类似符号，也有文献支撑，可以进行编码。但实际应用则提倡使用汉文标点符号系统，不再另外设计。 （二）数字符号 传统文献采用声、韵、调相拼的办法解决数字拼读问题，后来有关专家帮助设计了专门符号，规范了傣文数字表示形式。经
过讨论，大家认为可以保留数字符号并进行编码。同时，实际应用可使用阿拉伯数字，按傣语拼读。 （三）特殊符号 金平傣文中，a带韵尾时分长短，用3个特殊符号与字母组合来表示，方案中目前未进行编码，是否增加，如何编码由傣文专家进一步讨论研究决定。 （四）字体字型 目前草案中收录了现有古籍和实际应用中出现的字符，新设计的“小篆体”字体字型也基本固定，从使用情况看得到广大傣族群众普遍认可，可以照此申报。 四、关于申报程序 会议认为，提供讨论的方案总体全面、科学、细致、准确、可行。讨论中，大家充分发表了意见，提出了修改建议，对完善方案具有重要帮助。会议决定，以本次会议讨论通过的方案为中国方案。其他个人方案由中国电子技术标准化研究院积极争取吸收可取成分充实本方案，避免技术冲突和申报过程中不必要的人为障碍。会议明确，由本方案研制专家罗文斌老师根据会议讨论提出的意见对方案进行完善，报云南省民语委办公室对方案文字表述等进行把关，最后报中国电子技术标准化研究院提交国际标准化组织。 会议充分肯定金平县民族宗教局、金平县傣学会对金平傣文字符编码国际标准研制申报工作的关心和支持，充分肯定罗文斌老师等为研制申报方案作出的艰苦努力。会议要求有关方面要积	极配合，将工作落到实处，争取早日完成，为构建中华民族共有精神家园做出积极贡献！ 参加人员：陈壮 陶扬 熊玉有 自文清 谭玉婷 龚传洪 刀洁 李有明 余启勇 罗文斌 刀明春 刀明芬 刘琼 罗忠成 王明君 刘素芬 会议记录：谭玉婷

Figure 9 On September 10, 2024, the EAC of Honghe Prefecture and the EAC of Yunnan Province held a seminar meeting for the purpose of developing and submitting the proposal for the Tai Xaa character set, the minutes of the meeting have been made public in this document in accordance with the law

Source: Yunnan EAC

Summary Form

ISO/IEC JTC 1/SC 2/WG 2 PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646 Please fill all the sections A, B and C below. Please read Principles and Procedures Document (P&P) from unicode.org/wg2/principles.html for guidelines and details before filling this form. Please ensure you are using the latest Form from unicode.org/wg2/summaryform.html. See also unicode.org/roadmaps for latest Roadmaps.

A. Administrative

1. Title:	<i>Proposal to encode the Tai Xaa script in UCS</i>
2. Requester's name:	<i>China National Body</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Member body contribution</i>
4. Submission date:	<i>April 22, 2025</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<i>yes</i>
(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):	<i>yes</i>
Proposed name of script:	<i>Tai Xaa</i>
b. The proposal is for addition of character(s) to an existing block:	<i>no</i>
Name of the existing block:	
2. Number of characters in proposal:	<i>76</i>
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary ☒	B.1-Specialized (small collection)
C-Major extinct	B.2-Specialized (large collection)
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?	<i>yes</i>
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<i>yes</i>
b. Are the character shapes attached in a legible form suitable for review?	<i>yes</i>
5. Fonts related:	
a. Who will provide the appropriate font to the Project Editor of 10646 for publishing the standard?	<i>Chen Zhuang</i>
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>Chen Zhuang, chenzh-zhuang@163.com</i>
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<i>yes</i>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>yes</i>
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)?	<i>no</i>

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 behavior information such as line breaks, widths etc., Combining behavior, Spacing behavior, Directional behavior, Default Collation behavior, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at www.unicode.org for such information on other scripts. Also see Unicode Character Database (www.unicode.org/reports/tr44) 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?	<i>no</i>
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.)?	<i>yes</i>
If YES, with whom? <i>Dai people lived in Yunnan Province, PRC</i>	
If YES, available relevant documents: <i>this document</i>	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	<i>yes</i>
Reference: <i>this document</i>	
4. The context of use for the proposed characters (type of use; common or rare)	<i>common</i>
Reference: _____	
5. Are the proposed characters in current use by the user community?	<i>yes</i>
If YES, where? Reference: <i>Jinping Miao, Yao, and Dai Autonomous County, Honghe Hani and Yi Autonomous Prefecture, Yunnan Province, PRC.</i> <i>(中华人民共和国云南省红河哈尼族彝族自治州金平苗族瑶族傣族自治县)</i>	
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	<i>no</i>
If YES, is a rationale provided? _____	
If YES, reference: _____	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	<i>yes</i>
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	<i>no</i>
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?	<i>no</i>
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, or could be confused with, an existing character?	<i>yes</i>
If YES, is a rationale for its inclusion provided? <i>yes</i>	
If YES, reference: <i>this document, section 2.1 (relations with Tai Viet)</i>	
11. Does the proposal include use of combining characters and/or use of composite sequences?	<i>yes</i>
If YES, is a rationale for such use provided? <i>yes</i>	
If YES, reference: <i>combining characters are an inherent part of the script</i>	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? <i>no</i>	
If YES, reference: _____	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	<i>no</i>
If YES, describe in detail (include attachment if necessary): _____	

13. Does the proposal contain any Ideographic compatibility characters?	<i>no</i>
If YES, are the equivalent corresponding unified ideographic characters identified? _____	
If YES, reference: _____	