Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation
Международная организация по стандартизации

Doc Type: **Working Group Document**
Title: **A proposal to Encode Alternative New Tai Lue Digit 1**
Source: **M. Hosken**
Status: **Individual Contribution**
Action: **For adoption by JTC1/SC2/WG2**
Date: **2008-Jan-24**

1. Introduction. ISO 10646 already includes New Tai Lue in the range U+1980 .. U+19DF. As part of that encoding there is a NEW TAI LUE DIGIT ONE. New Tai Lue has two digit 1s. This proposal is to add the missing digit 1 to ISO 10646 and also to fix a problem with the current NEW TAI LUE DIGIT ONE.

In Tai Tham script on which New Tai Lue is based, there are two sets of digits. One of these sets was included into the New Tai Lue writing system, when it was invented, along with only the digit 1 coming from the other set. From a Tai Tham point of view, the set that was incorporated into the New Tai Lue writing system was the hora set and the lone digit 1 was taken from the tham set of digits.

The digit one from the complete set used by New Tai Lue is visually similar to U+19B1 NEW TAI LUE VOWEL SIGN AA. Due to the possibility of visual confusion, this character is only used where visual confusion cannot occur, for example as part of a multi digit number using digits from the same set. The lone digit 1 is only used where visual ambiguity may occur, for example when a digit 1 occurs on its own.

An examination of the set of digits encoded in the New Tai Lue block shows that the current glyph for U+19D1 NEW TAI LUE DIGIT ONE does not fit with the rest of the set but is actually the lone digit one originating from the tham set. It is necessary to encode both digit 1s and it would be helpful to resolve the confusion by changing the representative glyph of U+19D1 NEW TAI LUE DIGIT ONE to fit with the rest of that set and then to encode a new digit 1 which encodes the lone digit from the tham set.

For this reason, this proposal consists of two actions:

- Change the representative glyph for U+19D1 NEW TAI LUE DIGIT ONE to be the same as U+19B1 NEW TAI LUE VOWEL SIGN AA.

- Add a new code U+19DA NEW TAI LUE THAM DIGIT ONE with representative glyph being U+19D1 NEW TAI LUE DIGIT ONE as found in ISO 10646 as currently ammended.

Example 1 shows a scan from a page including all three uses of digit 1 found in New Tai Lue texts and their proposed encoding.

၁ ၠ ၆၁၇ၔၽၯၶၐၯၰ

1、您早！

ၰ 、 သၰၟၣၚၴၝၐၔၺ！

Example 1

2. Rationale. No mention of this alternative (and more commonly used) digit 1 was made in any of the proposals for this script. This proposal is to address an issue that arose during initial production implementation of this script.

An alternative approach to addressing this issue is to not change the current representative glyph for NEW TAI LUE DIGIT ONE and to borrow a DIGIT ONE from either the Myanmar or proposed Lanna block. But all such cross script unification is frought with implementation issues. It requires that implementations know about such cross script unifications and specifically address them. The cost of adding such a cross script linkage for one character is far greater than supporting this proposal.

Currently implementations are stalled due to the lack of the extra digit one. This provides an opportunity to make the representative glyph change on the existing digit one with next to no cost. Once the extra digit one is encoded implementations will be able to proceed.

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]

Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines**
and details before filling this form.
Please ensure you are using the latest Form from http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html
See also http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest** *Roadmaps*.

A. Administrative

1. **Title:**	***New Tai Lue Tham Digit One***
2. Requester's name:	*Martin Hosken*
3. Requester type (Member body/Liaison/Individual contribution):	*Individual contribution*
4. Submission date:	
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	*X*
(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):

        Proposed name of script:

    b. The proposal is for addition of character(s) to an existing block: *X*

        Name of the existing block: *New Tai Lue*

2. Number of characters in proposal: *1*

3. Proposed category (select one from below - see section 2.2 of P&P document):

    A-Contemporary   *X*   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?   *SIL*

    If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools

    used: *http://scripts.sil.org/DaiBannaSIL*

6. References:

    a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?   *no*

    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?   *National Language Commission, Kunming, China*
        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?   *no*
    Reference:

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

5. Are the proposed characters in current use by the user community?   *yes*
    If YES, where? Reference:

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: *addition to existing BMP block*

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?   *yes*
        If YES, is a rationale for its inclusion provided?   *yes*
        If YES, reference: *this document*

11. Does the proposal include use of combining characters and/or use of composite sequences?   *no*
    If YES, is a rationale for such use provided?
        If YES, reference:
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
        If YES, reference:

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: