Lao Tham in Terms of Lanna
A Response to L2/05-166 from L2/05-095

Martin Hosken,
SIL International and Payap University

Introduction

L2/05-095 is a proposal for the encoding of the Lanna Script. Around the same time as this proposal was being prepared, Gregory Kourilsky was completing his Master's Thesis on the computerization of Lao Tham. L2/05-166 is a published paper based on his work.

Lao Tham would generally be recognised as being an academic script rather than a script in general use today. It is of historic interest with a sizeable historic corpus behind it and as such needs consideration for encoding.

In this short paper, I hope to show to what degree the current Lanna proposal already covers nearly all of Lao Tham as presented in L2/05-166. It is not possible to complete the process due to lack of a full description of Lao Tham. But my conclusion will be that the current proposal is not in any way incompatible with Lao Tham and a future small proposal will be enough to add the missing characters and complete the encoding.

Clearly Lao Tham will need its own font which may have some Lao Tham specific behaviours in it, for example the two forms of U+1A80 U+1AA6.

Comparison

By the time Lao Tham has developed into the current modern scripts that are known collectively as Lanna, much of the strong distinction between Pali and Lao is lost. It is certainly possible to identify the sources of words by using such an analysis, but for the use of the script, for all intents and purposes the distinction is lost. Therefore, whereas the Lanna proposal does support all the characters listed in L2/05-166, they are not necessarily structured in the way that a Lao Tham academician may desire.

Page 4

The consonant chart is covered by the Lanna proposal. I am assuming that the final Liquid is U+1AA9.

page 5

The first list of characters is covered: U+1A98, U+1A9B, U+1A9D, U+1AA1, U+1AA2, U+1AAA. Notice that U+1AA2 is the same as ra and may be covered by U+1AAB instead. The second row is covered using sequences U+1AA8 U+1A80 followed by: U+1A87, U+1A97, U+1A9F, U+1AA0, U+1AA3, U+1AA4 respectively.

The second table shows varioius alternate subscript forms:

Character	Base	Pali subscript	Lao subscript
∞	U+1A81	U+1A80 U+1A81	U+1A80 U+1A81 / U+1AB1

Character	Base	Pali subscript	Lao subscript
ꪇ	U+1A87	U+1AC8	U+1AC8 / U+1A80 U+1A87
ꪗ	U+1A97	U+1A80 U+1A97	U+1A80 U+1A97 / U+1AA2
ꪣ	U+1AA3	U+1A80 U+1AA3	U+1A80 U+1AA3

Notice that in the case of U+1AA3 it is assumed that the contextual rule is sufficient. If not, then a variant selector may be proposed to address the problem.

Page 12

The final table on page 12 shows the various alternative subjoined forms that are required for Lao Tham.

+	*F000*	*F004*	*F01B*	*F01D*
F042	U+1A80 U+1A81	U+1AC8	U+1A80 U+1AA3	*
F043	U+1AB1	U+1A80 U+1A97	U+1A80 U+1AA3	U+1A80 U+1AA6

The final column shows that there is a shape that is not yet supported in the Lanna proposal. Whether such a shape is context dependent or would need particular marking is up to a full Lao Tham proposal.

Page 13

The table on this page shows various font variations that should not be encoded but be handled as font variants.

Conclusion

From L2/05-166 Lao Tham looks to fit within the existing Lanna proposal model. It is probable that some addition will need to be made to Lanna after further research into Lao Tham has been done. Since Lao Tham is an academic script rather than one that on its own justification would exist in the BMP, it is suggested that the Lanna proposal continue through the acceptance process and not be held up to wait for the further research into Lao Tham to be completed. This paper shows that the existing proposal will not hinder a later proposal or require a separate block.