

Thai/Lao collation weighting rules

Kent Karlsson
2003-06-03

Introduction

The collation rules for Thai, and probably also Lao, in the 14651 CTT (and the UTS 10 data files) do not follow the (informal) collation rules given in annex C.2,Thai string ordering, of ISO/IEC 14651. The following text is a quote from ISO/IEC 14651 beginning of annex C.2:

The widely accepted standard for Thai lexicographical ordering is defined in the Royal Institute Dictionary 2525 B.E. Edition (1982 A.D.), the official standard Thai dictionary. The ordering principles are:

- Words are ordered alphabetically, not phonetically. Consonants order is:

ກ ແ ຂ ຄ ຜ ພ ຂ ຈ ຊ ຊ ດ ດ ປ ປ ວ ວ ປ ປ ຖ ຖ ຕ ຕ ທ ທ ນ

ບ ປ ຜ ພ ພ ກ ມ ຍ ຮ ວ ອ ລ ກ ກ ວ ສ ຊ ສ ຫ ອ ສ

(ດ ດ ກ ກ ກ ກ are vowels and ligatures, but put in the order according to the sounds they represent.)

- Vowels are also ordered by written forms, not by sounds. Vowels order is:

ຂ ແ ຂ ແ ຂ ແ ຂ ແ ຂ ແ ຂ ແ ຂ

(ອ ອ ຍ are always ordered as consonants, although they sometimes act as vowels.)

String comparison is performed from left to right, but considering initial consonants before vowels in the same syllable.

- Tones and diacritics are ignored at level 1, at level 2 the order is:

ᬁ ጀ ጀ ጀ ጀ ጀ

In the following section, formal ordering rules that attempts to follow the informal description in that annex are given. Note that, even though Thai and Lao are Brahmic based scripts (Indic scripts), the Thai virama ([THAI CHARACTER PHINTHU](#)) appears to have lost its “conjoining” quality, and is not even handled as a silent vowel, but a mark ignored at level 1 of the collation weighting. This appears to be so because the virama is most often not written out explicitly, even though the inherent vowel is made silent in pronunciation. The resulting ambiguities (in pronunciation and semantics) have to be resolved by context. However it does not affect collation. Dependent vowels (and would-be dependent vowels) are still, in this proposal, ordered like dependent vowels for Indic scripts, at the very end of the weighting scale for level 1, along with Hangul vowels.

The rules below also contain a number of contractions. These are not explicitly declared below, even though that need be made for a complete 14651 style set of ordering rules; the contractions are easily inferred by the names of the contractions. Using these contractions as below, special prehandling for Thai and Lao in prehandling (which is now specified in UTS 10) is no longer needed (if it is decided to accept this part of this proposal). This method works also with ISO/IEC 14651, which does not require that there is any prehandling at all.

Suggestion for the CTT for Thai and Lao scripts

%%% Thai/Lao punctuation:

<U0E46> IGNORE; IGNORE; IGNORE; <U0E46>	% THAI CHARACTER MAIYAMOK	Lm repetition (not modifier letters, more)
<U0EC6> IGNORE; IGNORE; IGNORE; <U0EC6>	% LAO KO LA	Lm repetition (like apostrophe and hyphen)
<U0E4F> IGNORE; IGNORE; IGNORE; <U0E4F>	% THAI CHARACTER FONGMAN	Po bullet
<U0E2F> IGNORE; IGNORE; IGNORE; <U0E2F>	% THAI CHARACTER PAIYANNOI	Lo danda (not a letter!)
<U0EAF> IGNORE; IGNORE; IGNORE; <U0EAF>	% LAO ELLIPSIS	Lo danda (not a letter!)
<U0E5A> IGNORE; IGNORE; IGNORE; <U0E5A>	% THAI CHARACTER ANGKHANKHU	Po double danda
<U0E5A_U0E30> IGNORE; IGNORE; IGNORE; <U0E5A_U0E30>	% THAI CHARACTER ANGKHANKHU, THAI CHARACTER SARA A	end mark
<U0E5B> IGNORE; IGNORE; IGNORE; <U0E5B>	% THAI CHARACTER KHOMUT	Po end mark

%%% Thai/Lao tone marks and other diacritics. Placed after a consonant (*before* any dependent vowel?):

<U0E4E> IGNORE;<D0E4E>;<MIN>;<U0E4E>	% THAI CHARACTER YAMAKKAN	Mn	above
<U0E3A> IGNORE;<D0E3A>;<MIN>;<U0E3A>	% THAI CHARACTER PHINTHU	virama,	below, silent vowel
<U0E4C> IGNORE;<D0E4C>;<MIN>;<U0E4C>	% THAI CHARACTER THANTHAKHAT	Mn	above cancellation mark
<U0ECC> IGNORE;<D0E4C>;<MIN>;<U0ECC>	% LAO CANCELLATION MARK	Mn	above cancellation mark
<U0E47> IGNORE;<D0E47>;<MIN>;<U0E47>	% THAI CHARACTER MAITAIKHU	Mn	above (vowel?)
<U0E48> IGNORE;<D0E48>;<MIN>;<U0E48>	% THAI CHARACTER MAI EK	Mn	above
<U0EC8> IGNORE;<D0E48>;<MIN>;<U0EC8>	% LAO TONE MAI EK	Mn	above
<U0E49> IGNORE;<D0E49>;<MIN>;<U0E49>	% THAI CHARACTER MAI THO	Mn	above
<U0EC9> IGNORE;<D0E49>;<MIN>;<U0EC9>	% LAO TONE MAI THO	Mn	above
<U0E4A> IGNORE;<D0E4A>;<MIN>;<U0E4A>	% THAI CHARACTER MAI TRI	Mn	above
<U0ECA> IGNORE;<D0E4A>;<MIN>;<U0ECA>	% LAO TONE MAI TI	Mn	above
<U0E4B> IGNORE;<D0E4B>;<MIN>;<U0E4B>	% THAI CHARACTER MAI CHATTAWA	Mn	above
<U0ECB> IGNORE;<D0E4B>;<MIN>;<U0ECB>	% LAO TONE MAI CATAWA	Mn	above
<U0E4D> IGNORE;<D0E4D>;<MIN>;<U0E4D>	% THAI CHARACTER NIKHAHIT	Mn	above * final nasal
<U0ECD> IGNORE;<D0E4D>;<MIN>;<U0ECD>	% LAO NIGGAHITA	Mn	above * final nasal
<U0E45> IGNORE;<D0E45>;<MIN>;<U0E45>	% THAI CHARACTER LAKKHANGYAO	Lo	* length mark used with 0E24 or 0E26 (not a vowel!)

%%% Thai consonants:

<U0E01> <SOE01>;<BASE>;<MIN>;<U0E01>	% THAI CHARACTER KO KAI
<U0E40_U0E01> "<SOE01><SOE40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E01>	% E, KO KAI
<U0E41_U0E01> "<SOE01><SOE41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E01>	% AE, KO KAI
<U0E42_U0E01> "<SOE01><SOE42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E01>	% O, KO KAI
<U0E43_U0E01> "<SOE01><SOE43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E01>	% AI MAIMUAN, KO KAI
<U0E44_U0E01> "<SOE01><SOE44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E01>	% AI MAIMALAI, KO KAI
<U0E02> <SOE02>;<BASE>;<MIN>;<U0E02>	% THAI CHARACTER KHO KHAI
<U0E40_U0E02> "<SOE02><SOE40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E02>	% E, KHO KHAI
<U0E41_U0E02> "<SOE02><SOE41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E02>	% AE, KHO KHAI
<U0E42_U0E02> "<SOE02><SOE42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E02>	% O, KHO KHAI
<U0E43_U0E02> "<SOE02><SOE43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E02>	% AI MAIMUAN, KHO KHAI
<U0E44_U0E02> "<SOE02><SOE44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E02>	% AI MAIMALAI, KHO KHAI
<U0E03> <SOE03>;<BASE>;<MIN>;<U0E03>	% THAI CHARACTER KHO KHUAT

<U0E43_U0E17> "<SOE17><SOE43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E17> % AI MAIMUAN, THO THAHAN
 <U0E44_U0E17> "<SOE17><SOE44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E17> % AI MAIMALAI, THO THAHAN
 <U0E18> <SOE18>; <BASE>; <MIN>; <U0E18> % THAI CHARACTER THO THONG
 <U0E40_U0E18> "<SOE18><SOE40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E18> % E, THO THONG
 <U0E41_U0E18> "<SOE18><SOE41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E18> % AE, THO THONG
 <U0E42_U0E18> "<SOE18><SOE42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E18> % O, THO THONG
 <U0E43_U0E18> "<SOE18><SOE43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E18> % AI MAIMUAN, THO THONG
 <U0E44_U0E18> "<SOE18><SOE44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E18> % AI MAIMALAI, THO THONG
 <U0E19> <SOE19>; <BASE>; <MIN>; <U0E19> % THAI CHARACTER NO NU
 <U0E40_U0E19> "<SOE19><SOE40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E19> % E, NO NU
 <U0E41_U0E19> "<SOE19><SOE41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E19> % AE, NO NU
 <U0E42_U0E19> "<SOE19><SOE42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E19> % O, NO NU
 <U0E43_U0E19> "<SOE19><SOE43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E19> % AI MAIMUAN, NO NU
 <U0E44_U0E19> "<SOE19><SOE44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E19> % AI MAIMALAI, NO NU
 <U0E1A> <SOE1A>; <BASE>; <MIN>; <U0E1A> % THAI CHARACTER BO BAIMAI
 <U0E40_U0E1A> "<SOE1A><SOE40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E1A> % E, BO BAIMAI
 <U0E41_U0E1A> "<SOE1A><SOE41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E1A> % AE, BO BAIMAI
 <U0E42_U0E1A> "<SOE1A><SOE42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E1A> % O, BO BAIMAI
 <U0E43_U0E1A> "<SOE1A><SOE43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E1A> % AI MAIMUAN, BO BAIMAI
 <U0E44_U0E1A> "<SOE1A><SOE44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E1A> % AI MAIMALAI, BO BAIMAI
 <U0E1B> <SOE1B>; <BASE>; <MIN>; <U0E1B> % THAI CHARACTER PO PLA
 <U0E40_U0E1B> "<SOE1B><SOE40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E1B> % E, PO PLA
 <U0E41_U0E1B> "<SOE1B><SOE41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E1B> % AE, PO PLA
 <U0E42_U0E1B> "<SOE1B><SOE42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E1B> % O, PO PLA
 <U0E43_U0E1B> "<SOE1B><SOE43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E1B> % AI MAIMUAN, PO PLA
 <U0E44_U0E1B> "<SOE1B><SOE44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E1B> % AI MAIMALAI, PO PLA
 <U0E1C> <SOE1C>; <BASE>; <MIN>; <U0E1C> % THAI CHARACTER PHO PHUNG
 <U0E40_U0E1C> "<SOE1C><SOE40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E1C> % E, PHO PHUNG
 <U0E41_U0E1C> "<SOE1C><SOE41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E1C> % AE, PHO PHUNG
 <U0E42_U0E1C> "<SOE1C><SOE42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E1C> % O, PHO PHUNG
 <U0E43_U0E1C> "<SOE1C><SOE43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E1C> % AI MAIMUAN, PHO PHUNG
 <U0E44_U0E1C> "<SOE1C><SOE44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E1C> % AI MAIMALAI, PHO PHUNG
 <U0E1D> <SOE1D>; <BASE>; <MIN>; <U0E1D> % THAI CHARACTER FO FA
 <U0E40_U0E1D> "<SOE1D><SOE40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E1D> % E, FO FA
 <U0E41_U0E1D> "<SOE1D><SOE41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E1D> % AE, FO FA
 <U0E42_U0E1D> "<SOE1D><SOE42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E1D> % O, FO FA
 <U0E43_U0E1D> "<SOE1D><SOE43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E1D> % AI MAIMUAN, FO FA
 <U0E44_U0E1D> "<SOE1D><SOE44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E1D> % AI MAIMALAI, FO FA
 <U0E1E> <SOE1E>; <BASE>; <MIN>; <U0E1E> % THAI CHARACTER PHO PHAN
 <U0E40_U0E1E> "<SOE1E><SOE40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E1E> % E, PHO PHAN
 <U0E41_U0E1E> "<SOE1E><SOE41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E1E> % AE, PHO PHAN

<U0E42_U0E1E> "<S0E1E><S0E42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E1E> % O, PHO PHAN
 <U0E43_U0E1E> "<S0E1E><S0E43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E1E> % AI MAIMUAN, PHO PHAN
 <U0E44_U0E1E> "<S0E1E><S0E44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E1E> % AI MAIMALAI, PHO PHAN
 <U0E1F> <S0E1F>; <BASE>; <MIN>; <U0E1F> % THAI CHARACTER FO FAN
 <U0E40_U0E1F> "<S0E1F><S0E40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E1F> % E, FO FAN
 <U0E41_U0E1F> "<S0E1F><S0E41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E1F> % AE, FO FAN
 <U0E42_U0E1F> "<S0E1F><S0E42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E1F> % O, FO FAN
 <U0E43_U0E1F> "<S0E1F><S0E43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E1F> % AI MAIMUAN, FO FAN
 <U0E44_U0E1F> "<S0E1F><S0E44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E1F> % AI MAIMALAI, FO FAN
 <U0E20> <S0E20>; <BASE>; <MIN>; <U0E20> % THAI CHARACTER PHO SAMPHAO
 <U0E40_U0E20> "<S0E20><S0E40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E20> % E, PHO SAMPHAO
 <U0E41_U0E20> "<S0E20><S0E41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E20> % AE, PHO SAMPHAO
 <U0E42_U0E20> "<S0E20><S0E42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E20> % O, PHO SAMPHAO
 <U0E43_U0E20> "<S0E20><S0E43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E20> % AI MAIMUAN, PHO SAMPHAO
 <U0E44_U0E20> "<S0E20><S0E44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E20> % AI MAIMALAI, PHO SAMPHAO
 <U0E21> <S0E21>; <BASE>; <MIN>; <U0E21> % THAI CHARACTER MO MA
 <U0E40_U0E21> "<S0E21><S0E40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E21> % E, MO MA
 <U0E41_U0E21> "<S0E21><S0E41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E21> % AE, MO MA
 <U0E42_U0E21> "<S0E21><S0E42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E21> % O, MO MA
 <U0E43_U0E21> "<S0E21><S0E43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E21> % AI MAIMUAN, MO MA
 <U0E44_U0E21> "<S0E21><S0E44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E21> % AI MAIMALAI, MO MA
 <U0E22> <S0E22>; <BASE>; <MIN>; <U0E22> % THAI CHARACTER YO YAK
 <U0E40_U0E22> "<S0E22><S0E40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E22> % E, YO YAK
 <U0E41_U0E22> "<S0E22><S0E41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E22> % AE, YO YAK
 <U0E42_U0E22> "<S0E22><S0E42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E22> % O, YO YAK
 <U0E43_U0E22> "<S0E22><S0E43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E22> % AI MAIMUAN, YO YAK
 <U0E44_U0E22> "<S0E22><S0E44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E22> % AI MAIMALAI, YO YAK
 <U0E23> <S0E23>; <BASE>; <MIN>; <U0E23> % THAI CHARACTER RO RUA
 <U0E40_U0E23> "<S0E23><S0E40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E23> % E, RO RUA
 <U0E41_U0E23> "<S0E23><S0E41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E23> % AE, RO RUA
 <U0E42_U0E23> "<S0E23><S0E42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E23> % O, RO RUA
 <U0E43_U0E23> "<S0E23><S0E43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E23> % AI MAIMUAN, RO RUA
 <U0E44_U0E23> "<S0E23><S0E44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E23> % AI MAIMALAI, RO RUA
 <U0E24> <S0E24>; <BASE>; <MIN>; <U0E24> % THAI CHARACTER RU
 <U0E40_U0E24> "<S0E24><S0E40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E24> % E, RU
 <U0E41_U0E24> "<S0E24><S0E41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E24> % AE, RU
 <U0E42_U0E24> "<S0E24><S0E42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E24> % O, RU
 <U0E43_U0E24> "<S0E24><S0E43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E24> % AI MAIMUAN, RU
 <U0E44_U0E24> "<S0E24><S0E44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E24> % AI MAIMALAI, RU
 <U0E25> <S0E25>; <BASE>; <MIN>; <U0E25> % THAI CHARACTER LO LING
 <U0E40_U0E25> "<S0E25><S0E40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E25> % E, LO LING


```

<U0E40_U0E2C> "<S0E2C><S0E40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E2C> % E, LO CHULA
<U0E41_U0E2C> "<S0E2C><S0E41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E2C> % AE, LO CHULA
<U0E42_U0E2C> "<S0E2C><S0E42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E2C> % O, LO CHULA
<U0E43_U0E2C> "<S0E2C><S0E43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E2C> % AI MAIMUAN, LO CHULA
<U0E44_U0E2C> "<S0E2C><S0E44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E2C> % AI MAIMALAI, LO CHULA
<U0E2D> <S0E2D>; <BASE>; <MIN>; <U0E2D> % THAI CHARACTER O ANG
<U0E40_U0E2D> "<S0E2D><S0E40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E2D> % E, O ANG
<U0E41_U0E2D> "<S0E2D><S0E41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E2D> % AE, O ANG
<U0E42_U0E2D> "<S0E2D><S0E42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E2D> % O, O ANG
<U0E43_U0E2D> "<S0E2D><S0E43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E2D> % AI MAIMUAN, O ANG
<U0E44_U0E2D> "<S0E2D><S0E44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E2D> % AI MAIMALAI, O ANG
<U0E2E> <S0E2E>; <BASE>; <MIN>; <U0E2E> % THAI CHARACTER HO NOKHUK
<U0E40_U0E2E> "<S0E2E><S0E40>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E40_U0E2E> % E, HO NOKHUK
<U0E41_U0E2E> "<S0E2E><S0E41>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E41_U0E2E> % AE, HO NOKHUK
<U0E42_U0E2E> "<S0E2E><S0E42>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E42_U0E2E> % O, HO NOKHUK
<U0E43_U0E2E> "<S0E2E><S0E43>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E43_U0E2E> % AI MAIMUAN, HO NOKHUK
<U0E44_U0E2E> "<S0E2E><S0E44>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0E44_U0E2E> % AI MAIMALAI, HO NOKHUK

```

%%% Lao consonants:

```

<U0E81> <S0E81>; <BASE>; <MIN>; <U0E81> % LAO LETTER KO
<U0EC0_U0E81> "<S0E81><S0EC0>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC0_U0E81> % E, KO
<U0EC1_U0E81> "<S0E81><S0EC1>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC1_U0E81> % EI, KO
<U0EC2_U0E81> "<S0E81><S0EC2>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC2_U0E81> % O, KO
<U0EC3_U0E81> "<S0E81><S0EC3>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC3_U0E81> % AY, KO
<U0EC4_U0E81> "<S0E81><S0EC4>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC4_U0E81> % AI, KO
<U0E82> <S0E82>; <BASE>; <MIN>; <U0E82> % LAO LETTER KHO SUNG
<U0EC0_U0E82> "<S0E82><S0EC0>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC0_U0E82> % E, KHO SUNG
<U0EC1_U0E82> "<S0E82><S0EC1>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC1_U0E82> % EI, KHO SUNG
<U0EC2_U0E82> "<S0E82><S0EC2>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC2_U0E82> % O, KHO SUNG
<U0EC3_U0E82> "<S0E82><S0EC3>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC3_U0E82> % AY, KHO SUNG
<U0EC4_U0E82> "<S0E82><S0EC4>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC4_U0E82> % AI, KHO SUNG
<U0E84> <S0E84>; <BASE>; <MIN>; <U0E84> % LAO LETTER KHO TAM
<U0EC0_U0E84> "<S0E84><S0EC0>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC0_U0E84> % E, KHO TAM
<U0EC1_U0E84> "<S0E84><S0EC1>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC1_U0E84> % EI, KHO TAM
<U0EC2_U0E84> "<S0E84><S0EC2>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC2_U0E84> % O, KHO TAM
<U0EC3_U0E84> "<S0E84><S0EC3>"; "<BASE><BASE>"; "<MIN><MIN>"; <U0EC3_U0E84> % AY, KHO TAM

```



```

<U0EC1_U0EAA> "<S0EAA><S0EC1>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC1_U0EAA> % EI, SO SUNG
<U0EC2_U0EAA> "<S0EAA><S0EC2>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC2_U0EAA> % O, SO SUNG
<U0EC3_U0EAA> "<S0EAA><S0EC3>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC3_U0EAA> % AY, SO SUNG
<U0EC4_U0EAA> "<S0EAA><S0EC4>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC4_U0EAA> % AI, SO SUNG
<U0EAB> <S0EAB>;<BASE>;<MIN>;<U0EAB> % LAO LETTER HO SUNG
<U0EC0_U0EAB> "<S0EAB><S0EC0>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC0_U0EAB> % E, HO SUNG
<U0EC1_U0EAB> "<S0EAB><S0EC1>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC1_U0EAB> % EI, HO SUNG
<U0EC2_U0EAB> "<S0EAB><S0EC2>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC2_U0EAB> % O, HO SUNG
<U0EC3_U0EAB> "<S0EAB><S0EC3>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC3_U0EAB> % AY, HO SUNG
<U0EC4_U0EAB> "<S0EAB><S0EC4>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC4_U0EAB> % AI, HO SUNG

<U0EDC> "<S0EAB><S0E99>" ;<BASE><BASE>" ;<COMPAT><COMPAT>" ;<U0EDC> % LAO HO NO
<U0EC0_U0EDC> "<S0EAB><S0EC0><S0E99>" ;<BASE><BASE><BASE>" ;<COMPAT><COMPAT><COMPAT>" ;<U0EC0_U0EDC> % E, HO NO
<U0EC1_U0EDC> "<S0EAB><S0EC1><S0E99>" ;<BASE><BASE><BASE>" ;<COMPAT><COMPAT><COMPAT>" ;<U0EC1_U0EDC> % EI, HO NO
<U0EC2_U0EDC> "<S0EAB><S0EC2><S0E99>" ;<BASE><BASE><BASE>" ;<COMPAT><COMPAT><COMPAT>" ;<U0EC2_U0EDC> % O, HO NO
<U0EC3_U0EDC> "<S0EAB><S0EC3><S0E99>" ;<BASE><BASE><BASE>" ;<COMPAT><COMPAT><COMPAT>" ;<U0EC3_U0EDC> % AY, HO NO
<U0EC4_U0EDC> "<S0EAB><S0EC4><S0E99>" ;<BASE><BASE><BASE>" ;<COMPAT><COMPAT><COMPAT>" ;<U0EC3_U0EDC> % AI, HO NO
<U0EDD> "<S0EAB><S0EA1>" ;<BASE><BASE>" ;<COMPAT><COMPAT>" ;<U0EDD> % LAO HO MO
<U0EC0_U0EDD> "<S0EAB><S0EC0><S0EA1>" ;<BASE><BASE><BASE>" ;<COMPAT><COMPAT><COMPAT>" ;<U0EC0_U0EDD> % E, HO MO
<U0EC1_U0EDD> "<S0EAB><S0EC1><S0EA1>" ;<BASE><BASE><BASE>" ;<COMPAT><COMPAT><COMPAT>" ;<U0EC1_U0EDD> % EI, HO MO
<U0EC2_U0EDD> "<S0EAB><S0EC2><S0EA1>" ;<BASE><BASE><BASE>" ;<COMPAT><COMPAT><COMPAT>" ;<U0EC2_U0EDD> % O, HO MO
<U0EC3_U0EDD> "<S0EAB><S0EC3><S0EA1>" ;<BASE><BASE><BASE>" ;<COMPAT><COMPAT><COMPAT>" ;<U0EC3_U0EDD> % AY, HO MO
<U0EC4_U0EDD> "<S0EAB><S0EC4><S0EA1>" ;<BASE><BASE><BASE>" ;<COMPAT><COMPAT><COMPAT>" ;<U0EC4_U0EDD> % AI, HO MO

<U0EAD> <S0EAD>;<BASE>;<MIN>;<U0EAD> % LAO LETTER O
<U0EC0_U0EAD> "<S0EAD><S0EC0>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC0_U0EAD> % E, O
<U0EC1_U0EAD> "<S0EAD><S0EC1>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC1_U0EAD> % EI, O
<U0EC2_U0EAD> "<S0EAD><S0EC2>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC2_U0EAD> % O, O
<U0EC3_U0EAD> "<S0EAD><S0EC3>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC3_U0EAD> % AY, O
<U0EC4_U0EAD> "<S0EAD><S0EC4>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC4_U0EAD> % AI, O
<U0EAE> <S0EAE>;<BASE>;<MIN>;<U0EAE> % LAO LETTER HO TAM
<U0EC0_U0EAE> "<S0EAE><S0EC0>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC0_U0EAE> % E, HO TAM
<U0EC1_U0EAE> "<S0EAE><S0EC1>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC1_U0EAE> % EI, HO TAM
<U0EC2_U0EAE> "<S0EAE><S0EC2>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC2_U0EAE> % O, HO TAM
<U0EC3_U0EAE> "<S0EAE><S0EC3>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC3_U0EAE> % AY, HO TAM
<U0EC4_U0EAE> "<S0EAE><S0EC4>" ;<BASE><BASE>" ;<MIN><MIN>" ;<U0EC4_U0EAE> % AI, HO TAM

```

%%%%%%%%%%%%% After all scripts, among the Hangul vowels and other Indic vowels %%%

%%% Thai (dependent) vowels:

```
<U0E30> <S0E30>;<BASE>;<MIN>;<U0E30> % THAI CHARACTER SARA A
<U0E31> <S0E31>;<BASE>;<MIN>;<U0E31> % THAI CHARACTER MAI HAN-AKAT
<U0E32> <S0E32>;<BASE>;<MIN>;<U0E32> % THAI CHARACTER SARA AA
<U0E4D_U0E32> <S0E33>;<BASE>;<MIN>;<U0E33> % THAI CHARACTER NIKHAHIT, THAI CHARACTER SARA AA
<U0E33> <S0E33>;<BASE>;<COMPAT>;<U0E33> % THAI CHARACTER SARA AM
<U0E34> <S0E34>;<BASE>;<MIN>;<U0E34> % THAI CHARACTER SARA I
<U0E35> <S0E35>;<BASE>;<MIN>;<U0E35> % THAI CHARACTER SARA II
<U0E36> <S0E36>;<BASE>;<MIN>;<U0E36> % THAI CHARACTER SARA UE
<U0E37> <S0E37>;<BASE>;<MIN>;<U0E37> % THAI CHARACTER SARA UEE
<U0E38> <S0E38>;<BASE>;<MIN>;<U0E38> % THAI CHARACTER SARA U
<U0E39> <S0E39>;<BASE>;<MIN>;<U0E39> % THAI CHARACTER SARA UU
```

%%% Thai reordering (dependent) vowels:

```
<U0E40> <S0E40>;<BASE>;<MIN>;<U0E40> % THAI CHARACTER SARA E
<U0E41> <S0E41>;<BASE>;<MIN>;<U0E41> % THAI CHARACTER SARA AE
<U0E42> <S0E42>;<BASE>;<MIN>;<U0E42> % THAI CHARACTER SARA O
<U0E43> <S0E43>;<BASE>;<MIN>;<U0E43> % THAI CHARACTER SARA AI MAIMUAN
<U0E44> <S0E44>;<BASE>;<MIN>;<U0E44> % THAI CHARACTER SARA AI MAIMALAI
```

%%%%%%%%%%%%% Logical place is after some consonants(!), including tone mark or
%%%%%%%%%%%%% nikhahit applied to the consonants. Ordering pretends it is one vowel
%%%%%%%%%%%%% after the next consonant, while it may actually be pronounced even later.

%%% Lao (dependent) vowels:

```
<U0EB0> <S0EB0>;<BASE>;<MIN>;<U0EB0> % LAO VOWEL SIGN A
<U0EB1> <S0EB1>;<BASE>;<MIN>;<U0EB1> % LAO VOWEL SIGN MAI KAN
<U0EB2> <S0EB2>;<BASE>;<MIN>;<U0EB2> % LAO VOWEL SIGN AA
<U0ECD_U0EB2> <S0EB3>;<BASE>;<MIN>;<U0EB3> % LAO NIGGAHITA, LAO VOWEL SIGN AA
<U0EB3> <S0EB3>;<BASE>;<COMPAT>;<U0EB3> % LAO VOWEL SIGN AM
<U0EB4> <S0EB4>;<BASE>;<MIN>;<U0EB4> % LAO VOWEL SIGN I
<U0EB5> <S0EB5>;<BASE>;<MIN>;<U0EB5> % LAO VOWEL SIGN II
<U0EB6> <S0EB6>;<BASE>;<MIN>;<U0EB6> % LAO VOWEL SIGN Y
<U0EB7> <S0EB7>;<BASE>;<MIN>;<U0EB7> % LAO VOWEL SIGN YY
<U0EB8> <S0EB8>;<BASE>;<MIN>;<U0EB8> % LAO VOWEL SIGN U
<U0EB9> <S0EB9>;<BASE>;<MIN>;<U0EB9> % LAO VOWEL SIGN UU
<U0EBB> <S0EBB>;<BASE>;<MIN>;<U0EBB> % LAO VOWEL SIGN MAI KON
<U0EBC> <S0EBC>;<BASE>;<MIN>;<U0EBC> % LAO SEMIVOWEL SIGN LO
<U0EBD> <S0EBD>;<BASE>;<MIN>;<U0EBD> % LAO SEMIVOWEL SIGN NYO
```

%%% Lao reordering (dependent) vowels:

```
<U0EC0> <S0EC0>;<BASE>;<MIN>;<U0EC0> % LAO VOWEL SIGN E
<U0EC1> <S0EC1>;<BASE>;<MIN>;<U0EC1> % LAO VOWEL SIGN EI
<U0EC2> <S0EC2>;<BASE>;<MIN>;<U0EC2> % LAO VOWEL SIGN O
<U0EC3> <S0EC3>;<BASE>;<MIN>;<U0EC3> % LAO VOWEL SIGN AY
<U0EC4> <S0EC4>;<BASE>;<MIN>;<U0EC4> % LAO VOWEL SIGN AI
```

%%%%%%%%%%%%% Logical place is after some consonants(!), including tone mark or
%%%%%%%%%%%%% nikhahit applied to the consonants. Ordering pretends it is one vowel
%%%%%%%%%%%%% after the next consonant, while it may actually be pronounced even later.