The babel style for Vietnamese
Werner Lemberg, Han The Thanh, and Reinhard Kotucha
2015/12/31

1 Vietnamese language support

This file provides definitions for the Vietnamese language. It can be loaded by babel with either the vietnamese or the vietnam option. It works with xetex and luatex, as well as with 8-bit engines.

Unicode engines don’t require loading a font encoding, but you may use EU1 and EU2, respectively, if you want. In addition, you may load the language with the modifier licr to have access to LICR entities like \uhook, for example (but then fontenc must be loaded).

\usepackage[vietnamese.licr]{babel}

In 8-bit engines you must use the T5 font encoding, which always defines the LICR entities. If for some reason you don’t want the encoding to be loaded by the package, just use the modifier noencoding.

The code

\LdfInit{\CurrentOption}{captions\CurrentOption}

When this file is read as an option, i.e., by the \usepackage command, vietnamese could be an ‘unknown’ language in which case we have to make it known. So we check for the existence of \l@vietnamese and \l@vietnam to see whether we have to do something here.

\ifx\l@vietnamese@undefined
\ifx\l@vietnam@undefined
\nopatterns{Vietnamese}
\adddialect\l@vietnamese 0
\adddialect\l@vietnam 0
\else
\chardef\l@vietnamese\l@vietnam
\fi
\else
\ifx\l@vietnam@undefined
\chardef\l@vietnam\l@vietnamese\l@vietnam
\fi
\fi
\chardef\l@vietnam\l@vietnamese\l@vietnamese
If `utf8x` (from `ucs`) is used instead of `utf8`, do some additional adjustments.

```latex
\@ifpackageloaded{ucs}{%
  \PreloadUnicodePage{0}%
  \PreloadUnicodePage{1}%
  \PreloadUnicodePage{30}%
  \ifx\UnicodeCharFilter\@undefined
  \UCSProtectionIeC
  \else
  \UnicodeCharFilter\IeC
  \fi}{}

In 8-bit \LaTeX\ we load and set the encoding, but not yet (this temporary command is redefined for Unicode engines).
\def.vn@tempa{\
\@expandtwoargs\in@{,noencoding,}{,\BabelModifiers,}\
\ifin@\else
\InputIfFileExists{t5enc.def}\
{\message{Loading definitions for the Vietnamese font encoding}}\
{\errhelp{I can't find the file `t5enc.def' for Vietnamese fonts}}\
{\errormessage{Since I do not know what the T5 encoding means^^J
I can't typeset Vietnamese.^^J
I stop here, while you install the VnTeX package^^J
to get a suitable `t5enc.def' file}}\
\@@end}\
\@ifpackageloaded{inputenc}{}\
\PackageWarning{babel}{No input encoding specified for
Vietnamese}}%
\fi

\DeclareRobustCommand{\viettext}{\fontencoding{T5}\selectfont}
\let\viet=\viettext
\DeclareTextFontCommand{\textviet}{\viet}
```

The macro `\captionsvietnamese` defines all strings used in the four standard document classes provided with \LaTeX.  

\StartBabelCommands*{vietnamese,vietnam}{captions}
\[\textnormal{unicode, charset=utf8, fontenc=EU1 EU2}\]
\SetString{\prefacename}{Li nói u}
\SetString{\refname}{Tịi liu}
\SetString{\abstractname}{Tùm tt ni dung}
\SetString{\bibname}{Tịi liu tham kho}
\SetString{\chaptername}{Chng}
\SetString{\appendixname}{Ph lc}
\SetString{\contentname}{Mc lc}
\SetString{\listfigurename}{Danh sách hơnh v}
\SetString{\listtablename}{Danh sách bng}

\captionsvietnamese
And now the generic branch, using the LICR and assuming T5.

\StartBabelCommands{vietnamese,vietnam}{captions}
\SetString{prefacename}{L‘\`ohorn i n‘\oi \dj\`a d‘acircumflex u}%
\SetString{refname}{T‘ai li d‘ecircumflex u}%
\SetString{abstractname}{T‘\‘om t‘\‘abreve t n‘d‘ocircumflex i dung}%
\SetString{bibname}{T‘ai li d‘ecircumflex u tham kh‘h(a)\o}%
\SetString{chaptername}{Ch\‘uhorn \‘ohorn ng}%
\SetString{appendixname}{Ph\d{d(u)\l d(u)c}c}\%
\SetString{contentsname}{M\d{d(u)c \l d(u)c}c}\%
\SetString{listfigurename}{Danh s‘\‘ach h\‘\inh v\‘e}o}\%
\SetString{listtablename}{Danh s‘\‘ach b\‘h(a)ng}\%
\SetString{indexname}{Ch\‘hi m\d{d(u)c}c}\%
\SetString{figurename}{H\‘\inh}\%
\SetString{tablename}{B\‘h(a)ng}\%
\SetString{partname}{Ph\‘\‘acircumflex n}\%
\SetString{pagename}{Trang}\%
\SetString{headpagename}{Trang}\%
\SetString{seename}{Xem}\%
\SetString{alsoname}{Xem th\‘ecircumflex m}\%
\SetString{enclname}{K‘\‘em theo}\%
\SetString{ccname}{C‘ung gi\‘u}\%
\SetString{headtoname}{Gi}\%
\SetString{pagename}{Trang}\%
\SetString{figurename}{H\‘inh}\%
\SetString{tablename}{B\‘ang}\%
\SetString{partname}{Ph\‘\‘acircumflex n}\%
\SetString{enclname}{K‘\‘em theo}\%
\SetString{ccname}{C‘ung gi\‘u}\%
\SetString{pagename}{Trang}\%
\SetString{figurename}{H\‘inh}\%
\SetString{tablename}{B\‘ang}\%
\SetString{partname}{Ph\‘\‘acircumflex n}\%
The macro \extrasvietnamese performs all the extra definitions needed for the Vietnamese language. The macro \noextrasvietnamese is used to cancel the actions of \extrasvietnamese.

We specify that the Vietnamese group of shorthands should be used. These characters are ‘turned on’ once; later on their definition may vary. For Vietnamese texts \frenchspacing should be in effect. We make sure this is the case and reset it if necessary.

We restore the encoding with \latintext, but this is far from ideal. A better approach would be preferable.

2 LICR for Unicode

The file uvnenc.def maps the LICR to the actual characters if the Unicode engine is used.
A fake accent for the horn.

If the dblaccent package is not loaded, make a dummy \DeclareTextDoubleComposite.
\DeclareTextComposite{\d}{\vn@enc}{I}{7882} % Idotbelow
\DeclareTextComposite{\'}{\vn@enc}{O}{217} % Ograve
\DeclareTextComposite{\'}{\vn@enc}{O}{218} % Oacute
\DeclareTextComposite{\~}{\vn@enc}{O}{7910} % Otilde
\DeclareTextComposite{\h}{\vn@enc}{O}{7886} % Ohookabove
\DeclareTextComposite{\d}{\vn@enc}{O}{7884} % Odotbelow
\DeclareTextComposite{\`}{\vn@enc}{O}{360} % Ocircumflex
\DeclareTextDoubleComposite{\`}{\vn@enc}{\horn}{O}{7900} % Ohorngrave
\DeclareTextDoubleComposite{\'}{\vn@enc}{\horn}{O}{7898} % Ohornacute
\DeclareTextDoubleComposite{\~}{\vn@enc}{\horn}{O}{7904} % Ohorntilde
\DeclareTextDoubleComposite{\h}{\vn@enc}{\horn}{O}{7902} % Ohornhookabove
\DeclareTextDoubleComposite{\d}{\vn@enc}{\horn}{O}{7906} % Ohorndotbelow
\DeclareTextComposite{\horn}{\vn@enc}{U}{416} % Ohorn
\DeclareTextDoubleComposite{\`}{\vn@enc}{\horn}{U}{7914} % Uhorngrave
\DeclareTextDoubleComposite{\'}{\vn@enc}{\horn}{U}{7912} % Uhornacute
\DeclareTextDoubleComposite{\~}{\vn@enc}{\horn}{U}{7916} % Uhorntilde
\DeclareTextDoubleComposite{\h}{\vn@enc}{\horn}{U}{7920} % Uhornhookabove
\DeclareTextDoubleComposite{\d}{\vn@enc}{\horn}{U}{7924} % Uhorndotbelow
\DeclareTextComposite{\horn}{\vn@enc}{Y}{7922} % Yhorn
\DeclareTextDoubleComposite{\~}{\vn@enc}{Y}{928} % Ytilde
\DeclareTextComposite{\~}{\vn@enc}{Y}{928} % Ytilde
\DeclareTextComposite{\h}{\vn@enc}{Y}{7928} % Yhookabove
\DeclareTextComposite{\d}{\vn@enc}{Y}{7932} % Ydotbelow
\DeclareTextComposite{\`}{\vn@enc}{a}{7883} % agrave
\DeclareTextComposite{\'}{\vn@enc}{a}{242} % aacute
\DeclareTextComposite{\~}{\vn@enc}{a}{245} % atilde
\DeclareTextComposite{\h}{\vn@enc}{a}{7843} % ahookabove
\DeclareTextComposite{\d}{\vn@enc}{a}{7841} % adotbelow
\DeclareTextComposite{\`}{\vn@enc}{a}{243} % acircumflex
\DeclareTextComposite{\~}{\vn@enc}{\~}{a}{7847} % acircumflexgrave
\DeclareTextComposite{\h}{\vn@enc}{\~}{a}{7851} % acircumflexabove
\DeclareTextComposite{\d}{\vn@enc}{\Ecircumflex}{7878} % Ecircumflexdotbelow
\DeclareTextSymbol{\Ocircumflex}{\vn@enc}{360} % Ocircumflex
\DeclareTextComposite{\'}{\vn@enc}{\Ocircumflex}{7888} % Ocircumflexacute
\DeclareTextComposite{\~}{\vn@enc}{\Ocircumflex}{7894} % Ocircumflextilde
\DeclareTextComposite{\h}{\vn@enc}{\Ocircumflex}{7892} % Ocircumflexhookabove
\DeclareTextComposite{\d}{\vn@enc}{\Ocircumflex}{7896} % Ocircumflexdotbelow
\DeclareTextSymbol{\Ohorn}{\vn@enc}{416} % Ohorn
\DeclareTextComposite{\'}{\vn@enc}{\Ohorn}{7900} % Ohornacute
\DeclareTextComposite{\~}{\vn@enc}{\Ohorn}{7904} % Ohortilde
\DeclareTextComposite{\h}{\vn@enc}{\Ohorn}{7902} % Ohornhookabove
\DeclareTextComposite{\d}{\vn@enc}{\Ohorn}{7906} % Ohordotbelow
\DeclareTextSymbol{\acircumflex}{\vn@enc}{243} % acircumflex
\DeclareTextComposite{\'}{\vn@enc}{\acircumflex}{7845} % acircumflexacute
\DeclareTextComposite{\~}{\vn@enc}{\acircumflex}{7851} % acircumflextilde
\DeclareTextComposite{\h}{\vn@enc}{\acircumflex}{7849} % acircumflexhookabove
\DeclareTextComposite{\d}{\vn@enc}{\acircumflex}{7853} % acircumflexdotbelow
\DeclareTextSymbol{\abreve}{\vn@enc}{259} % abreve
\DeclareTextComposite{\'}{\vn@enc}{\abreve}{7855} % abreveacute
\DeclareTextComposite{\~}{\vn@enc}{\abreve}{7861} % abrevedotbelow
\DeclareTextComposite{\d}{\vn@enc}{\abreve}{7859} % abrevedotbelow
\DeclareTextSymbol{\ecircumflex}{\vn@enc}{7893} % ecircumflex
\DeclareTextComposite{\'}{\vn@enc}{\ecircumflex}{7873} % ecircumflexgrave
\DeclareTextComposite{\~}{\vn@enc}{\ecircumflex}{7877} % ecircumflextilde
\DeclareTextComposite{\h}{\vn@enc}{\ecircumflex}{7875} % ecircumflexhookabove
\DeclareTextComposite{\d}{\vn@enc}{\ecircumflex}{7879} % ecircumflexdotbelow
\DeclareTextSymbol{\ocircumflex}{\vn@enc}{7891} % ocircumflex
\DeclareTextComposite{\'}{\vn@enc}{\ocircumflex}{7889} % ocircumflexacute
\DeclareTextComposite{\~}{\vn@enc}{\ocircumflex}{7895} % ocircumflextilde
\DeclareTextComposite{\h}{\vn@enc}{\ocircumflex}{7893} % ocircumflexhookabove
\DeclareTextComposite{\d}{\vn@enc}{\ocircumflex}{7897} % ocircumflexdotbelow
\DeclareTextSymbol{\ohorn}{\vn@enc}{417} % ohorn
\DeclareTextComposite{'}{\vn@enc}{\ohorn}{7901} % ohorngrave
\DeclareTextComposite{\'}{\vn@enc}{\ohorn}{7899} % ohornacute
\DeclareTextComposite{\~}{\vn@enc}{\ohorn}{7905} % ohorntilde
\DeclareTextComposite{\h}{\vn@enc}{\ohorn}{7903} % ohornhookabove
\DeclareTextComposite{\d}{\vn@enc}{\ohorn}{7907} % ohorndotbelow
\DeclareTextSymbol{\uhorn}{\vn@enc}{432} % uhorn
\DeclareTextComposite{'}{\vn@enc}{\uhorn}{7915} % uhorngrave
\DeclareTextComposite{\'}{\vn@enc}{\uhorn}{7913} % uhornacute
\DeclareTextComposite{\~}{\vn@enc}{\uhorn}{7919} % uhorntilde
\DeclareTextComposite{\h}{\vn@enc}{\uhorn}{7917} % uhornhookabove
\DeclareTextComposite{\d}{\vn@enc}{\uhorn}{7921} % uhorndotbelow
\% Make these ugly names still valid (needed for use with utf8).
\DeclareTextSymbol{\ACIRCUMFLEX}{\vn@enc}{194} % Acircumflex
\DeclareTextComposite{'}{\vn@enc}{\ACIRCUMFLEX}{7846} % Acircumflexgrave
\DeclareTextComposite{\'}{\vn@enc}{\ACIRCUMFLEX}{7844} % Acircumflexacute
\DeclareTextComposite{\~}{\vn@enc}{\ACIRCUMFLEX}{7850} % Acircumflextilde
\DeclareTextComposite{\h}{\vn@enc}{\ACIRCUMFLEX}{7848} % Acircumflexhookabove
\DeclareTextComposite{\d}{\vn@enc}{\ACIRCUMFLEX}{7852} % Acircumflexdotbelow
\DeclareTextSymbol{\ABREVE}{\vn@enc}{258} % Abreve
\DeclareTextComposite{'}{\vn@enc}{\ABREVE}{7856} % Abrevegrave
\DeclareTextComposite{\'}{\vn@enc}{\ABREVE}{7854} % Abreveacute
\DeclareTextComposite{\~}{\vn@enc}{\ABREVE}{7860} % Abreetildetilde
\DeclareTextComposite{\h}{\vn@enc}{\ABREVE}{7858} % Abreehookabove
\DeclareTextComposite{\d}{\vn@enc}{\ABREVE}{7862} % Abreedotbelow
\DeclareTextSymbol{\ECIRCUMFLEX}{\vn@enc}{202} % Ecircumflex
\DeclareTextComposite{'}{\vn@enc}{\ECIRCUMFLEX}{7872} % Ecircumflexgrave
\DeclareTextComposite{\'}{\vn@enc}{\ECIRCUMFLEX}{7870} % Ecircumflexacute
\DeclareTextComposite{\~}{\vn@enc}{\ECIRCUMFLEX}{7876} % Ecircumflextilde
\DeclareTextComposite{\h}{\vn@enc}{\ECIRCUMFLEX}{7874} % Ecircumflexhookabove
\DeclareTextComposite{\d}{\vn@enc}{\ECIRCUMFLEX}{7878} % Ecircumflexdotbelow
\DeclareTextSymbol{\OCIRCUMFLEX}{\vn@enc}{360} % Ocircumflex
\DeclareTextComposite{'}{\vn@enc}{\OCIRCUMFLEX}{7890} % Ocircumflexgrave
\DeclareTextComposite{\'}{\vn@enc}{\OCIRCUMFLEX}{7888} % Ocircumflexacute
\DeclareTextComposite{\~}{\vn@enc}{\OCIRCUMFLEX}{7894} % Ocircumflextilde
\DeclareTextComposite{\h}{\vn@enc}{\OCIRCUMFLEX}{7892} % Ocircumflexhookabove
\DeclareTextComposite{\d}{\vn@enc}{\OCIRCUMFLEX}{7896} % Ocircumflexdotbelow
\DeclareTextSymbol{\OHORN}{\vn@enc}{416} % Ohorn
\DeclareTextComposite{'}{\vn@enc}{\OHORN}{7900} % Ohorngrave
\DeclareTextComposite{\'}{\vn@enc}{\OHORN}{7898} % Ohornacute
\DeclareTextComposite{\~}{\vn@enc}{\OHORN}{7904} % Ohorntilde
\DeclareTextComposite{\h}{\vn@enc}{\OHORN}{7902} % Ohornhookabove
Finally, we extend \@uclclist to make \MakeUppercase and \MakeLowercase work correctly. This case mapping is “global” and directly related to the encoding and the LICR, so we don’t use \SetCase, which is intended for small readjustments required by a certain language.

\expandafter\def\expandafter\@uclclist\expandafter{% 
\abreve\Abreve
\abreve\ABREVE
\acircumflex\Acircumflex
\acircumflex\ACIRCUMFLEX
\dj\DJ
\ecircumflex\Ecircumflex
\ecircumflex\ECIRCUMFLEX
\ocircumflex\Ocircumflex
\ocircumflex\OCIRCUMFLEX
\ohorn\Ohorn
\ohorn\OHORN
\uhorn\Uhorn
\uhorn\UHORN
\endinput

(/unilicr)