

Printing Vietnamese characters by adding diacritical marks via T_EX

Brother Eric Vogel FSC

Introduction

The technique described here uses T_EX to produce the diacritical marks for the various vowels needed for Vietnamese. The marks used are: accent grave, accent acute, tilde, question mark and dot (below). New vowels are introduced using the above by placing a hat (above a or e) or breve accent (above a) or by attaching a "beard" (using the breve accent) to o and u. The macros for these definitions are listed below in Table 3.

These macros were developed using PC T_EX; they will work equally well with any other T_EX. Other facilities described here, however, are specific to PC T_EX and to PC-Write.

Method of producing text with diacritical marks attached

The author of the Vietnamese text enters the text without the marks. The author then edits the text by placing the cursor on the letter needing a diacritical mark and striking a key sequence that produces the mark(s). What appears on the screen is the command for T_EX to reproduce the correct marks upon printing. For example to produce the Vietnamese word *Mag* the author starts with *Mag* on the screen and the cursor under the *a*. The author then presses control-k. On the screen appears `\gu`, followed by the original *a* and *g*. When the document is T_EXed, *Mag* is printed.

Definitions used by T_EX

The accents are produced by placing the correct command (i.e. definition) in front of the letter to which one wishes to attach the accent. Because of the additional space needed for o-beard and u-beard, these commands (definitions) contain the needed o or u (O or U); hence if one has text to which accents are being added, it is necessary to delete the o or u (O or U) which is in the text. The author has used PC-Write (Shareware) to define the introduction of these commands using either the control key, the shift key or control-c with an additional letter. Table 2 displays the letters as accented by this program as well as what one types using the editor of PC-Write and what is displayed on the screen. To define these control sequences it is necessary to add to the `ed.def` file of PC-Write the command `!viet.ext` in order that the editor can access the file `viet.ext` which is given in Table 1.

Special attention needed to be paid to producing the diacritical mark of a question mark over letters as the Vietnamese mark does not have a dot under the rest of the question mark. This was accomplished by using METAFONT to generate a new ? in the font `vfont`. Two files are needed and must be copied into appropriate subdirectories of PC T_EX: `vfont.tfm`, which is copied into the subdirectory `\textfms`, and `vfont.pk`, which is copied into the subdirectory `\pixel\dpi300`. The following code was used in naming the definitions: `a` indicates accent acute; `g` indicates accent grave; `t` indicates tilde; `q` indicates question mark; `d` indicates dot (under the vowel); `h` indicates hat; combinations of the above indicate more than one accent: e.g. `ah` indicates accent acute over hat. The accents that include a "beard" attached to either o or u all have a `b` in their code. `ob` indicates an o with a beard; `ub` indicates a u with a beard; `b` can be combined with the other accents given above: e.g. `ba` indicates a beard and an accent acute. The capitals for U and O have special commands (necessitated by the lack of keys on the keyboard). These take advantage of the special two stroke capabilities of PC-Write and are addressed by using control-c followed by a letter. The coding of the commands consists in using `co` to indicate a capital O and `cu` to indicate a capital U (with, of course `b` for beard) and the possibility of another letter for other accents. For example, `\cuba` indicates a capital U, with a beard, with accent acute. Because the author ran out of keys to be redefined using the above method before he ran out of needed keys, two additional keys were redefined: `@` and `#` as shown below.

Obtaining the files

Upon your request, the author will send you a disk with four files: `vdefq` and `viet.ext` and the two fonts for printing: `vfont.tfm` and `vfont.pk`. The first contains all the macros used to produce the definitions of the diacritical marks via T_EX. The second is the file needed by PC-Write (to be placed in the `ed.def` file of that software) to define the control sequences which will produce the correct results. Of course the macros can be used with any software that is capable of defining enough keys. The two fonts for printing must be copied into the correct subdirectories as was described above. (Any donations gladly accepted.)

Brother Eric Vogel FSC
P. O. Box 5150
Saint Mary's College
Moraga, Calif. 94575

Table 1. Key redefining used by PC-Write. Parentheses not included. Found in file: viet.ext.

064:092,113,117,032 (Redefines @ key to be \qu, question over breve)
 035:092,117,100,032 (Redefines # key to be \ud, breve over, dot under)
 c:555, "CAPBEAR"; (Beginning of redefinition of keys via two strokes)
 556,113, "\cob "; (Ctrl-c, q. Capital O with a beard)
 556,119, "\cobg "; (Ctrl-c, w. Capital O with a beard and accent grave)
 556,101, "\coba "; (Ctrl-c, e. Capital O with a beard and accent acute)
 556,114, "\cobt "; (Ctrl-c, r. Capital O with a beard and tilde)
 556,116, "\cobq "; (Ctrl-c, t. Capital O with a beard and question mark)
 556,121, "\codb "; (Ctrl-c, y. Capital O with a beard and dot below)
 556,097, "\cub "; (Ctrl-c, a. Capital U with a beard)
 556,115, "\cubg "; (Ctrl-c, s. Capital U with a beard and accent grave)
 556,100, "\cuba "; (Ctrl-c, d. Capital U with a beard and accent acute)
 556,102, "\cubt "; (Ctrl-c, f. Capital U with a beard and tilde)
 556,103, "\cubq "; (Ctrl-c, g. Capital U with a beard and question mark)
 556,104, "\cubd " (Ctrl-c, h. Capital U with a beard and dot below)
 z: "\g " (Ctrl-z. Accent grave)
 x: "\a " (Ctrl-x. Accent acute)
 v: "\t " (Ctrl-v. Tilde)
 b: "\q " (Ctrl-b. Question mark)
 n: "\d " (Ctrl-n. Dot under letter)
 m: "\tu " (Ctrl-m. Tilde over breve)
 a: "\ub " (Ctrl-m. Lower case u with a beard)
 s: "\bg u\s " (Ctrl-s. u with a beard and accent grave)
 d: "\ba u\s " (Ctrl-d. u with a beard and accent acute)
 f: "\bt u\s " (Ctrl-f. u with a beard and tilde)
 g: "\bq u\s " (Ctrl-g. u with a beard and question mark)
 h: "\bd u\s " (Ctrl-h. u with a beard and dot below)
 j: "\u " (Ctrl-j. Accent breve)
 k: "\gu " (Ctrl-k. Accent grave over breve)
 l: "\au " (Ctrl-l. Accent acute over breve)
 q: "\ob " (Ctrl-q. Lower case o with a beard)
 w: "\bg o\s " (Ctrl-w. o with beard and accent grave)
 e: "\ba o\s " (Ctrl-e. o with beard and accent acute)
 r: "\bt o\s " (Ctrl-r. o with beard and tilde)
 t: "\bq o\s " (Ctrl-t. o with beard and question mark)
 y: "\bd o\s " (Ctrl-y. o with a beard and dot below)
 u: "\h " (Ctrl-u. Hat over letter)
 i: "\gh " (Ctrl-i. Accent grave over hat)
 o: "\ah " (Ctrl-o. Accent acute over hat)
 p: "\th " (Ctrl-p. Tilde over hat)
 [: "\qh " (Ctrl-[. Question mark over hat)
]: "\hd " (Ctrl-]. Hat over, dot under)

Table 2. Table showing how diacritical marks appear, are produced and the commands that appear on the screen

	VOWEL	GRAVE	ACUTE	TILDE	QUEST.	DOT
	A	À	Á	Ã	Â	Ạ
	a	à	á	ã	â	ạ
TYPE		ctl-z	ctl-x	ctl-v	ctl-b	ctl-n
SCREEN		\g	\a	\t	\q	\d
	Â	Â	Á	Ã	Â	Ạ
	â	â	á	ã	â	ạ
TYPE	ctl-u	ctl-i	ctl-o	ctl-p	ctl-[ctl-]
SCREEN	\h	\gh	\ah	\th	\qh	\hd
	Ă	Ă	Ă	Ă	Ă	Ă
	ă	ă	ă	ă	ă	ă
TYPE	ctl-j	ctl-k	ctl-l	ctl-m	shift-2	shift-3
SCREEN	\u	\gu	\au	\tu	\qu	\ud
	E	È	É	Ê	Ê	Ệ
	e	è	é	ê	ê	ệ
TYPE		ctl-z	ctl-x	ctl-v	ctl-b	ctl-n
SCREEN		\g	\a	\t	\q	\d
	Ê	Ê	É	Ê	Ê	Ệ
	ê	ê	é	ê	ê	ệ
TYPE	ctl-u	ctl-i	ctl-o	ctl-p	ctl-[ctl-]
SCREEN	\h	\gh	\ah	\th	\qh	\hd
	I	Ì	Í	Ĩ	Ỉ	Ị
	i	ì	í	ĩ	ỉ	ị
TYPE		ctl-z	ctl-x	ctl-v	ctl-b	ctl-n
SCREEN		\g	\a	\t	\q	\d

	Ò	Ò	Ó	Õ	Ö	Ï
	ó	ò	ó	õ	ö	ï
TYPE		ctl-z	ctl-x	ctl-v	ctl-b	ctl-n
SCREEN		\g	\a	\t	\q	\d
	Ô	Ô	Ó	Õ	Ö	Ï
	ô	ò	ó	õ	ö	ï
TYPE	ctl-u	ctl-i	ctl-o	ctl-p	ctl-[ctl-]
SCREEN	\h	\gh	\ah	\th	\qh	\hd
	U	Û	Ú	Û	Û	Û
	u	ù	ú	û	ü	ÿ
TYPE		ctl-z	ctl-x	ctl-v	ctl-b	ctl-n
SCREEN		\g	\a	\t	\q	\d
	ø	ø	ó	õ	ö	ø
TYPE	ctl-q	ctl-w *	ctl-e *	ctl-r *	ctl-t *	ctl-y *
SCREEN	\ob	\bg o\s	\ba o\s	\bt o\s	\bq o\s	\bd o\s
	ur	ür	úr	ûr	ÿr	ÿr
TYPE	ctl-a	ctl-s *	ctl-d *	ctl-f *	ctl-g *	ctl-h *
SCREEN	\ub	\bg u\s	\ba u\s	\bt u\s	\bq u\s	\bd u\s
	Œ	Œ	Œ	Œ	Œ	Œ
TYPE	ctl-c q *	ctl-c w *	ctl-c e *	ctl-c r *	ctl-c t *	ctl-c y *
SCREEN	\cob	\cobg	\coba	\cobt	\cobq	\cobd
	Ÿ	Û	Ú	Û	Û	Û
TYPE	ctl-c a *	ctl-c s *	ctl-c d *	ctl-c f *	ctl-c g *	ctl-c h *
SCREEN	\cub	\cubg	\cuba	\cubt	\cubq	\cubd

* indicates that the control sequence introduces its own letter and hence the letter present in the text must be deleted.

Table 3. Macros for definitions used. Found in file: vdefq.

```

\font\v=vfont
\def\g{\'}
\def\af{\'}
\def\t{\~}
\def\h{\^}
\def\au{\leavevmode\raise.929ex\rlap{\kern.1em\accent"13 }%
\raise.465ex\rlap{\kern.1em\accent"15 }}
\def\hd{\leavevmode\raise.465ex\rlap{\kern.1em\accent"5E }%
\lower.465ex\rlap{\kern.2em.}}
\def\q{\leavevmode\raise1.394ex\rlap{\kern.2em\v ?}}
\def\gu{\leavevmode\raise.929ex\rlap{\kern.1em\accent"12 }%
\raise.465ex\rlap{\kern.1em\accent"15 }}
\def\tu{\leavevmode\raise1.161ex\rlap{\kern.1em\accent"7E }%
\raise.465ex\rlap{\kern.1em\accent"15 }}
\def\qu{\leavevmode\raise1.858ex\rlap{\kern.2em\v ?}}%
\raise.465ex\rlap{\kern.1em\accent"15 }}
\def\ud{\leavevmode\raise.465ex\rlap{\kern.1em\accent"15 }%
\lower.465ex\rlap{\kern.15em.}}
\def\gh{\leavevmode\raise1.161ex\rlap{\kern.15em\accent"12 }%
\raise.465ex\rlap{\kern.1em\accent"5E }}
\def\ah{\leavevmode\raise1.161ex\rlap{\kern.05em\accent"13 }%
\raise.465ex\rlap{\kern.1em\accent"5E }}
\def\th{\leavevmode\raise1.161ex\rlap{\kern.1em\accent"7E }%
\raise.465ex\rlap{\kern.1em\accent"5E }}
\def\qh{\leavevmode\raise2.090ex\rlap{\kern.2em\v ?}}%
\raise.465ex\rlap{\kern.1em\accent"5E }}
\def\ob{o\lower.604ex\hbox{\kern-.25em\accent"15 }} % o with a beard
\def\ub{u\lower.604ex\hbox{\kern-.25em\accent"15 }} % u with a beard
\def\s{\hbox{\kern.2em}} % adds two pts space
\def\bg#1{\'#1\b}
\def\ba#1{\'#1\b}
\def\bt#1{\~#1\b}
\def\bq#1{\q#1\b}
\def\bd#1{\d#1\b}
\def\b{\lower.604ex\hbox{\kern-.25em\accent"15 }}
\def\cob{O\b} % O with a beard
\def\cub{U\b} % U with a beard
\def\cobg{\leavevmode\raise.698ex\rlap{\kern.2em\accent"12 }\cob}
% O-beard and accent grave
\def\coba{\leavevmode\raise.698ex\rlap{\kern.2em\accent"13 }\cob}
% O-beard and accent acute
\def\cobt{\leavevmode\raise.698ex\rlap{\kern.2em\accent"7E }\cob}
% O-beard and title
\def\cobd{\leavevmode\lower.465ex\rlap{\kern.2em.}\cob}
% O-beard and dot below
\def\cobq{\leavevmode\raise1.626ex\rlap{\kern.3em\v ?}\cob}
% O-beard and question mark
\def\cubg{\leavevmode\raise.698ex\rlap{\kern.2em\accent"12 }\cub}
% O-beard and accent grave
\def\cuba{\leavevmode\raise.698ex\rlap{\kern.2em\accent"13 }\cub}
% O-beard and accent acute
\def\cubt{\leavevmode\raise.465ex\rlap{\kern.15em\accent"7E }\cub}
% O-beard and title
\def\cubq{\leavevmode\raise1.626ex\rlap{\kern.25em\v ?}\cub}
% U-beard and question mark
\def\cubd{\leavevmode\lower.465ex\rlap{\kern.2em.}\cub}
% U-beard and dot below

```