Hellenic Polytonic HOWTO

Table of Contents

Introduction

The current state of handling Hellenic language characters on GNU / Linux systems is examined. Emphasis is placed on the use of Hellenic polytonic (accented) fonts and keyboard drivers.

Try the pdf version of this document if the html version is not rendered correctly.

History

Information update – more free fonts 20060505 - tlgu.carmen.gr
Information update – X information 20060422
Information update – gnome information 20051011
Date correction – information request 20050830
Amendments, addition of keyboard layouts 20050722
First release: 20050717 – cap8.gr/tlgu

Acknowledgements

Sindi Keesan, for her valuable notes on the first draft.
Simos Xenitellis, for his “gr” keyboard and font updates.

References

bash documentation (man bash) - GNU Bourne-Again Shell, Free Software Foundation, Inc., 2002 (8-bit input and output), other shell program manuals, as referenced in the text

The Linux keyboard and console HOWTO, Andries Brouwer, 12-Oct-2002

X keyboard configuration, www.x-docs.org/XKB/XKBproto.pdf, www.xfree86.org/current/XKB-Config.html, www.xfree86.org/current/XKB-Enhancing.html

Unicode consortium code sets, www.unicode.org

The abolition of the polytonic system - www.kairatos.com.gr/polytoniko.htm
(A brief history of the accented script, in Hellenic)

Accentuation Tutorial ist-socrates.berkeley.edu/~ancgreek/accenthtml/accentuation.html

Polytonic keyboard use guide, Simos Xenitellis and Costas Pistiolis (in Hellenic), planet.hellug.gr/misc/polytonic/WritePolytonic.pdf

Multi-lingual text on Linux, Jan Willem Stumpel, www.jw-stumpel.nl/stestu

Greek Font Society, www.greekfontsociety.org/pages/en_typefaces.html

Disclaimer

The usual disclaimer about misconfiguring your system beyond repair or oblueterating your work applies: Don't blame it on me. Do one thing at a time. Read and Learn. Write to tlgu, carmen gr, in case this document contains inaccuracies, errors or if you have some information that others can benefit from.

Accents and “Breathings”

All ancient (a few centuries BCE - “Before Current Era”) Hellenic writings I have seen use what we now call upper-case or **capital letters** exclusively (both modern Hellenic and “Latin” forms). Words were not even separated from one another. Many times, a scribe would write left-to-right and continue right-to-left (“voustrofedon” - the way oxen move when plowing a field). People learned and knew how to pronounce each word.

Lower-case letters (actually called “**peza” – prosaic**) appeared during the “Hellenistic” times, when the Hellenic language was the prevailing one in many parts of the world. **Aristophanes Byzantius** around 200 BCE separated the words and introduced lower-case letters as a tool for indicating the **correct pronunciation** of ancient script – as opposed to the “common” language – to language learners (a diction guide) *[need some documentation-reference on this one]*. Indeed, lower case letters reflect the form assumed by mouth, tongue and cheeks for voicing each letter.

Accent and breathing marks survived several centuries and were eventually abolished from official use in 1982 save for the **tonos** (accent) and **dialytika** (diaeresis). Like in other parts of the world, “capital” or “city” pronunciation had taken over at one point in time, rendering the original spelling and accentuation of the words obsolete. Fortunately, the correct (as in older - original) pronunciation of many Hellenic words is still preserved in local Hellenic dialects and other languages, as well.

I will provide a couple of **examples**, hoping that I am not far off the mark. The well-known beauty abducted by the Trojan prince is now called “Elenee”, the first two “e”s pronounced like the second “e” in el-**e**-ctron. This word used to start with a “rough breathing” mark (daseia) which may be pronounced by making the first consonant “denser” (Helen, Yelena) but also as a “c”, like in “super” or “hyper” (same rough breathing in front) which would give us Celene (Celine) or Celena (not surprisingly, this is the most common name for earth's lone and bright satellite in Hellenic). As if this was not confusing enough, the current spelling of this word's root is “selen-” (Selene, selenium), with an “s” sound rather than a “z” sound.

Note also the last vowel – an “eta” - which is interchangeably pronounced as a long “ee” or as an “a”.

Same thing with John. You may hear it pronounced “E-o-annees” in Hellenic but the rough breathing in front of the first vowel indicates that the original pronunciation would be much closer to “Yannis” or “Johannes”.

Back to the subject – the most common of these **accents** are the **acute** (oxeia), the **grave** (vareia or bareia) and the **circumflex** (perispomeni – distinct from surrounding voicings). The **diaeresis** (dialytica) shows that two vowels are pronounced separately, rather than as a dipthong. There is a **iota subscript** (hypogegrammeni or prosgegrammeni, depending on position) that may indicate a short “e” (iota) after the subscripted vowel. There are also marks to indicate long and short (**macron** and **breve - brachy**).

There are two “breathings”, **smooth** (psili – thin) and **rough** (daseia - dense) that may co-exist with the other accents.

The Hellenic word for accent (stress) is “**tonos**”. The term **polytonic** refers to a script utilizing pronunciation marks (stressed syllable and optionally change in tone and/or separation of stressed syllable, pronunciation modifiers), where the plain (contemporary or monotonic) Hellenic script only uses one accent to indicate the syllable that must be stressed.

- Yes, but how were the ancient hellenic words actually pronounced?

Lacking the necessary expertise I can only propose the following, to be pursued outside this HOWTO: Look in your local dialects for words that their insides (stems) look suspiciously close to known ancient accented words. Let us collect them and correlate them with both accented and unaccented forms. Try not to be influenced by established “schools” on this matter. See where it leads us. Example: ῥοή – βροχή, ὁρῶ – θορῶ.

Output

Computers use output devices, like display screens and printers, to communicate stored information to the outside world (for the purpose of this discussion, the “outside world” is you, the computer user).

Information is stored in digital computers as a sequence of binary digits (bits). Therefore, a number of schemes have been devised to represent written language symbols in computer memory, as needs and technology change. We call these codification schemes **internal representation**.

These stored symbols are eventually translated into a graphical (written) representation to be displayed on a computer screen or to be engraved on a more permanent medium. These graphical representations are called **glyphs** (carving strokes). Different **fonts** (assortments of character types) have different glyphs for the same internal representation.

Internal symbol representation may be translated to glyphs in a variety of ways. In

the older days of typewriters, when one wanted to type an accented character, one would press the respective accent symbol key(s). The typewriter platen would not move until a letter symbol was typed. This allowed the creation of a large number of glyphs. You could use any accent symbol over (or under) any letter character on your keyboard.

On the other hand, real typographers have typographical elements with all (usual) accented character combinations already cast to facilitate their strenuous job of arranging small pieces of metalwork (type) into their letterpress type holder.

Two different methods, both implemented in today's display and font processing computer programs. Using the latter method, ready-made accented character glyphs are available in our computer's type boxes, selectable by a **unique code** (one-to-one representation). Using the former “dead accent” method, a **sequence of symbol codes** is translated into a letter-plus-accent(s) combination glyph to be displayed or printed.

Both methods have their pros and cons. We may use our computers to “set type” and communicate our thoughts to others but we also use them to process and re-arrange data in an effort to extract useful information.

Input

The usual method of entering textual information in a computer at this time is by pressing keyboard buttons (keys). This key-press information is eventually translated into internal computer representation codes and stored in computer memory.

Again, the model of our computer's keyboard is the keyboard of a typewriter: letter keys (capitals), number and symbol keys. To save on machinery each typewriter key actuated a stem holding more than one metal-type symbol (two or even three characters).

Which means that some mechanism was needed to select one or other type element. The Shift key would lower or raise the platen relative to the type-holders. Later came the IBM Selectric that translated the (by then) customary shift-key combinations to a movement of a type-ball bringing the desired type element before the platen, followed by the daisy-wheel (a daisy-like structure with flexible radial elements, carrying the type). Type-balls or daisy-wheels had to match the keyboard layout, however.

User-programmable computers allow more flexibility. Today's keyboards output specific digital codes when a key is pressed (and released) reflecting the key's position on the keyboard. In turn, this **keyboard code** is translated into **internal representation** depending on the **keyboard translation** used and the current state of **keyboard modifiers**. This last one is easier to understand. A “Shift” key or a “Capitals Lock” key or a “Control” key is a keyboard modifier. Meaning that the state of the particular key or combination (e.g. the Control and Alt keys

pressed at the same time) affects the internal representation code ultimately produced by keyboard interpretation (translation) programs.

The same mechanism may be extended to accent key presses. When a key representing an accent is pressed, this is remembered until a key representing a letter is pressed. The combined effect of accent and letter key presses may be translated into a single code – the internal representation of a complex character.

A great benefit, reflecting on the price of keyboards over the years, is that hardware manufacturers don't have to produce different keyboards for different languages. Keyboard electronics stay the same and the only thing that may need to be changed is cheap plastic key caps.

Internal representation

A great number of communications and computer code schemes have been designed and implemented to represent written-language symbols over the years. Some of them are still used though permitting very limited (for current standards) character sets, like the Baudot-Murray and ASCII code sets (32 and 128 codes, respectively).

But the word is multi-lingual and extensions to these limited code sets were devised to allow for multi-lingual character sets in computers where the basic storage element is 8 bits wide (256 code combinations). The ISO 8859 family of code sets – still in use today – represents such an extension, where the plain Latin alphabet is represented within the first 128 different codes and the remaining 128 codes have struggled to carry our world heritage.

Obviously, the large number of distinct accented Hellenic characters could not fit in such a limited space. A solution would be the adoption of “combining marks” to form distinct glyphs (at processing / display time). But then there were Asian language symbols, ancient language symbols, artificial language symbols, musical symbols, mathematical / engineering symbols – humans are an inventive lot. And, guess what, each one wants to have access to symbols of special interest along with the usual language symbols. And we do have programmable computers, right?

Enter the Universal Coded Character Set (UCS) efforts, also known as ISO 10646 or **Unicode**, aiming at merging the symbols of all different cultures, current and older, into a unified scheme. This scheme defines both single-point complex characters, as well as multi-character combinations.

Of interest to this discussion are the non-accented and single-accent Hellenic characters in the **Unicode Greek and Coptic** code set (codes numbered 0370-03FF), the accents and accented character combinations in the **Greek Extended** code set (1F00-1FFF), as well as the **Combining Diacritical Marks** code set (0300-036F). These codes are adequate for representing the largest part of Hellenic literature, dating back a few thousand years.

More symbols are constantly being proposed for adoption by Unicode like ancient mathematical and engineering notation, ancient and medieval musical notation and even older Hellenic language symbols.

GNU/Linux keyboard translation

Our favorite operating system kernel is quite flexible regarding the translation of keyboard codes into different kinds of internal representation, permitting a relatively high-level definition of how key presses and key press combinations should be handled.

On the other hand, graphical user interfaces (GUI) like the X Window System and window managers are known to add their own layer of keyboard translations and code handling.

Multiple keyboard layouts are allowed. By entering a user-defined key combination on your keyboard you are able to switch between different keyboard layouts. In addition, you may select different sets of keyboard layouts for each application or window.

Hellenic (polytonic) under X

Almost all GNU / Linux distributions about the size of a CD-ROM or more carry a full complement of keyboard translation tables and several Unicode fonts, including the basic (single-accent) Hellenic glyphs. To get accented Hellenic characters on your display you need to install **fonts** that include glyphs for the code sets mentioned above – at least.

If you want to produce accented Hellenic script you also need to install the proper flavor of the Hellenic **x-keyboard translation table**.

Hellenic polytonic fonts

Some quite elegant fonts have seen the light in recent years, including Hellenic accented characters and much more. Search the web for “unicode fonts polytonic Hellenic Greek” to find sites for downloading **Gentium** by Victor Gaultney, **Cardo** by David Perry, **Galatia SIL** by Jonathan Kew, **Galilee Unicode Gk** by Rodney J. Decker, **Athena Unicode** by Jeffrey Rusten - Ralph Hancock, **Porson** by Richard Spaulding, Greek Font Society (GFS) fonts: **GFS Didot** and **GFS Bodoni** by Takis Katsoulidis and George Matthiopoulos, **GFS Neohellenic** (rendered from the original Scholderer “New Hellenic”), **GFS Olga** (rendered from the original Porson font - Hellenic characters only), **DejaVu Serif** by Stepan Roh, Ben Laenen, **MgOpenCanonica** by Magenta Ltd. and several others. All the above are fonts freely available. There are also several non-free fonts that you may elect to use.

Font installation may be performed manually (read the Font - HOWTO) or handled

by system utility programs, like KDE's **Control Center**, or by software package managers, such as Red Hat's Package Manager (**rpm**), Debian's **apt-get** or **synaptic** or graphical front-ends like **kpackage**, or SuSE's **YaST**.

Under KDE, one may click on the "**K**" virtual button and then select Control Center > System Administration > Font Installer. Make sure you use "Administrator Mode" for system-wide font availability. There is also a built-in font previewer.

In addition, other font browsers may be available in your systems, like **xfontmap** and **tkfont**. Also, newer font browser / character map utilities (see No polytonic keyboard later in this text).

Hellenic polytonic keyboard under X

The X-server will translate **keycodes** representing the physical keys to **keysyms** based on the contents of the key **modifier list** and the **keymap table**. Keyboard symbol codes are arranged in **groups**. A user may switch between groups using a special code. Hmmm... sounds a bit complicated but let's see an actual example.

For **on-the-fly** keyboard changes, the **setxkbmap** program will evoke the appropriate X-server calls. But we must also keep in mind that keyboard translations are affected by **locale** environment data or user "language and cultural preferences". What we are interested in here, is a single environment variable (LC_CTYPE) and the commands below may be placed in a file called **.profile** (in the user's home directory) for automatic execution after logging-in. A discussion on **locales** will follow.

```
export LC_CTYPE=el_GR.UTF-8
setxkbmap us,gr -variant ,polytonic -option grp:ctrl_shift_toggle -option
grp_led:scroll
```

or

```
export LC_CTYPE=el_GR.UTF-8
setxkbmap us,el -variant ,polytonic -option grp:ctrl_shift_toggle -option
grp_led:scroll
```

[Note the change of the language code from** (h)**el**(lenic) to **gr**(eek) **in 2005; apart from philosophical implications – language code versus territorial code – breaks backward compatibility and discourages people – links are cheap in GNU/Linux – please use them!]

The first line will set the **locale** variable affecting character handling (LC_CTYPE) to Hellenic, Unicode (UTF-8).

The second line is the one doing the actual loading of the keyboard translation. This one defines two translation groups: U.S. English and Hellenic – polytonic variant. Now, by pressing Ctrl/Shift, the keyboard layout will change from the plain Latin one to Hellenic, allowing the creation of Unicode internal

representation corresponding to the key pressed.

Leave out the word **polytonic** and you are left with a contemporary Hellenic (single-accent) keyboard.

Also, check out this one:

```
setxkbmap us,gr,ru -option grp:alt_shift_toggle,grp_led:scroll
```

or

```
setxkbmap us,el,ru -option grp:alt_shift_toggle,grp_led:scroll
```

In addition to the U.S. English, Hellenic and Russian keyboards, you also have the Scroll Lock LED to indicate that your keyboard is not in the default group. Also note that group shifts are now effected by Alt/Shift rather than Ctrl/Shift.

If you keep playing with options, you may find that you do not get the expected behavior. To clear stored options, use the command without any option arguments (**man setxkbmap**):

```
setxkbmap -option
```

To set the default **system-wide** keyboard behavior, the system administrator (root) may edit the /etc/X11/XF86Config file (after saving a backup copy...)

Mine reads:

```
Section "InputDevice"
  Driver       "Keyboard"
  Identifier   "Keyboard[0]"
  Option       "MapName" "Standard Keyboard [ pc104 ]"
  Option       "Protocol" "Standard"
  Option       "XkbLayout" "us"
  Option       "XkbModel" "pc104"
  Option       "XkbRules" "xfree86"
EndSection
```

I could modify it to read as follows, in order to be able to change groups using Ctrl/Shift and have a visual indication on my Scroll Lock LED:

```
Section "InputDevice"
  Driver       "Keyboard"
  Identifier   "Keyboard[0]"
  Option       "MapName" "Standard Keyboard [ pc104 ]"
  Option       "Protocol" "Standard"
  Option       "XkbLayout" "us"
  Option       "XkbModel" "pc104"
  Option       "XkbRules" "xfree86"
# use us,gr or us,el depending on whether your system supports one or the
other
```

```
#  Option       "XkbLayout" "us,el"
  Option       "XkbLayout" "us,gr"
  Option       "XkbVariant" ",polytonic"
  Option       "XKbOptions" "grp:ctrl_shift_toggle,grp_led:scroll"
EndSection
```

Note that the X-server needs to be restarted for these changes to take effect (Ctrl/Alt/Backspace)

My Hellenic keyboard description (**el**) is dated 3-Jun-2003, by Vasilis Vasaitis and Kostas Georgiou. There are newer ones, implementing the **gr** layout. Either or both may be found in

`/usr/X11R6/lib/X11/xkb/symbols;` the respective compose table may be found at `/usr/X11R6/lib/X11/locale/el_GR.UTF-8/Compose`

To see a graphical representation of the keyboard layout active in the current display, use the following commands:

```
xkbprint -nkg 2 -eps -lc el_GR.utf8 $DISPLAY kbd.ps
gv kbd.ps (or ggv kbd.ps)
```

There is also a **virtual keyboard** utility under KDE called **viki,** by Shaheed Haque, that will show your keyboard in real time (but not all interesting parts yet – ver. 3.4.0).

Let me try to depict then, the hellenic and “polytonic” variant keyboard layouts:

The Hellenic Polytonic keyboard:

This is the “**el polytonic** or **gr polytonic**” Hellenic layout, found in current (2005) distributions:

Tab	· : ;	Σ ς	€ Ε ε	Ρ ρ	Τ τ	Υ υ	Θ θ	Ι ι	Ο ο	Π π	‒ ¨ ~	˘ ‘
Caps Lock	Α α	Σ σ	Δ δ	Φ φ	Γ γ	Η η	Ξ ξ	Κ κ	Λ λ	’ ΄	‘ `	| \
Left Shift	Ζ ζ	Χ χ	Ψ ψ	Ω ω	Β β	Ν ν	Μ μ	< ,	> .	? /	Right Shift	

The two unshifted keys next to Lambda (L) are the dead-accent keys for the **acute** (oxeia) and **grave** (vareia). The same two keys, shifted, will produce psili and daseia.

The unshifted keys next to Pi (P) will produce the dead-accent keys **circumflex** (perispomeni) and hypogegrammeni – prosgegrammeni (iota **subscript**). Shifted, the key next to P will produce the dead-accent code for **diaeresis** (dialytika).

The same two keys, with the Right-Alt (Alt Gr) pressed will produce the dead-accent codes for **macron** and **breve** (brachy).

The “Q” key will produce the usual semi-colon and colon symbols, while with the Right-Alt (Alt Gr) pressed will produce the code for middle dot (used as semicolon – **ano teleia**).

The “E” key, with the Right-Alt (Alt Gr) pressed will produce the code for Euro { suspiciously looks like the combined sign for two drachmas and half an obol ;-) }

The Hellenic polytonic keyboard (**gr polytonic** in post-2005 xorg distributions), may look thus ***(untested – may change)***:

Tab	ʹ ῾	Σ ς	€ Ε ε	Ρ Ϙ ρ ϙ	Τ τ	Υ υ	Θ θ	Ι ι	Ο ο	Π ℼ π ℿ	‘ ~	˘ ¯
Caps Lock	Α α	Σ Ϛ σ ς	Δ δ	Φ Ϝ φ ϝ	Γ γ j	Η η	Ξ ξ	Κ ϗ κ ϟ	Λ λ	¨ ΄	’ ’	| `
Left Shift	Ζ ζ	Χ χ	Ψ ψ	Ω ω	Β β	Ν ν	Μ μ	< ,	> .	? /	Right Shift	

The unshifted key next to Lambda (L) is the dead-accent key for the **acute** (oxeia). The same key, shifted, will produce the dead-accent code for **diaeresis** (dialytika).

The next key to the right (' on your latin keyboard) is the dead-accent key for **psili**. The same key, shifted, will produce **daseia**.

Further to the right (\ on your latin keyboard) is the dead-accent key for **grave** (vareia).

The unshifted key next to Pi (P) will produce the dead-accent keys **circumflex** (perispomeni); the same key, shifted, will produce hypogegrammeni – prosgegrammeni (iota **subscript**).

The next key to the right will produce the dead-accent codes for **macron** (unshifted) and **breve** (brachy).

Following are codes produced by pressing the **acute** (; on your Latin keyboard) and another key (shifted or unshifted):

The code for middle dot (used as semicolon – **ano teleia**) is produced by **acute** followed by **period** (.)

The codes for **sampi** are produced by pressing **acute**, followed by **P**.

The codes for **sigmoid koppa** are produced by pressing **acute**, followed by **K**.

The codes for **koppa** are produced by pressing **acute**, followed by **R**.

The codes for **stigma** are produced by pressing **acute**, followed by **S**.

The codes for **digamma** are produced by pressing **acute**, followed by **F**.

The code for yot is produced by pressing **acute**, followed by **G** (shifted or unshifted).

The code for **dexia keraia** (top-right-to-lower-left numeral sign) is produced by pressing **acute**, followed by **q** (unshifted), while the code for **aristeri keraia** (lower-left-to-top-right) is produced by pressing **acute**, followed by **Q** (shifted) – Caps Lock will not work in this case.

The "standard" (monotonic) Hellenic keyboard:

Tab	· : ;	Σ ς	€ E ε	P ρ	T τ	Y υ	Θ θ	I ι	O ο	Π π	{ [	}]
Caps Lock	A α	Σ σ	Δ δ	Φ φ	Γ γ	H η	Ξ ξ	K κ	Λ λ	¨ ΄	" '	| \
Left Shift	Z ζ	X χ	Ψ ψ	Ω ω	B β	N ν	M μ	< ,	> .	? /	Right Shift	

Similar to the above, the dead accents are next to Lambda (L), limited to the "**tonos**" (stress mark) and diaeresis (**dialytica**).

Hellenic polytonic keyboard under KDE and GNOME

(May 2006 - THIS DOES NOT WORK FOR ALL X APPLICATIONS – USE THE INSTRUCTIONS IN THE PREVIOUS SECTION IF IT DOES NOT WORK FOR YOU)

This is the **KDE Control Center** way:

Click on the "**K**" virtual button on your KDE panel, Control Center > Regional & Accessibility > Keyboard Layout.

Check the "**Enable Keyboard Layouts**" box and select the **el (Greek)** layout.

Click on the “**Variant**” drop-down list and select “**polytonic**”, if you so wish.

You may also select whether the keyboard layout change is Global (all windows), separate for each Application or separate for each Window.

The default layout-switching key combination is... **Alt/Ctrl/K** (you guessed it, didn't you). Alas, this default combination is ill-fated as the moment you enter the layout switching combination, “K” is no longer “kay”, it is “kappa”. You are stuck with a Hellenic keyboard. You have to go back with the mouse and reconfigure.

Of course, you may use the next Control Center selection, called “**Keyboard Shortcuts**”. There, under the “**Keyboard**” section is the “**Switch to Next Keyboard Layout**” entry. Set this to something neutral, like Ctrl/Space or Alt/Enter. Or you may wish to try the “Multi-key” option, which lets you define combinations like Ctrl/Shift...

And this is the **Gnome Preferences** way:

On your Gnome Panel click **Desktop** and select **Gnome Control Center (Desktop Preferences)**, **Keyboard**.

Click on the **Layouts** tab and **Add Greek el(polytonic)** or **gr(polytonic)** – there is a variant selection triangle next to the Greek label. A nice feature is the ability to have a **Separate group for each window**, selectable by ticking the appropriate field in this tab. You can type Hellenic in one word processor window and English in another, without having to switch keyboard layouts.

You may then click on the **Layout Options** tab to define the **Group Shift/Lock behavior**. e.g. you may select **Control+Shift changes group**.

gswitchit is an applet to display a country flag on the panel (right click on the panel, select Add to Panel, Utilities...)

Locales

“A locale is a set of language and cultural rules. These cover aspects such as language for messages, different character sets, lexigraphic conventions, etc. A program needs to be able to determine its locale and act accordingly to be portable to different cultures.” (see **man 7 locale, man 1 locale**)

The locale setting (comprising a number of environment variables) may affect the way our keyboard works and the way our text is sorted and displayed. These variables may have already been set based on your (or someone else's) selections at system installation time. Also be aware of the fact that **file names may be encoded according to your locale setting**, rendering them inaccessible to other computer systems. Be very conservative on file names – use plain Latin characters as much as possible.

A locale name is of the form language[_territory][.codeset][@modifier]. A few examples:

```
LC_CTYPE=el_GR.UTF-8
LC_CTYPE=el_GR
LC_CTYPE=el_GR.ISO-8859-7
LC_CTYPE=el_GR.UTF-8@euro
```

For a list of all locales currently **supported** by your computer, use

```
locale -a
```

The default storage location for the locale files is: /usr/X11R6/lib/X11/locale

Let's take a look at some of these variables:

LC_ALL - sets all LC_xxx variables to a given locale (affects character translation, message translation, number representation, time format, collating sequence and monetary representation).

LC_CTYPE - determines what characters are allowed, and whether they should be treated as letters, digits, punctuation, or control characters.

LANG - determines the language used

The **first non-empty value** among these three determines the current locale for character handling, and in particular the **default text encoding**.

You may query current locale settings. e.g.

```
locale -k LC_CTYPE
```

Hellenic keyboard at the console (CLI)

To get Hellenic characters at the console (this is not a polytonic solution) we need the following:

A keyboard map.

A font with Hellenic characters.

We also need to tell the shell (bash) not to mutilate the codes produced by our keyboard translation table and the terminal subsystem to use the display codes loaded.

```
# load the Hellenic translation table
loadkeys gr

# load a corresponding font and unicode console translation map
setfont iso07.16
```

```
setfont -m 8859-7

# leave my characters alone
set input-meta on
set convert-meta off
set output-meta on

# switch on the user-defined character sets
echo -ne "\033(K\033)K" >/dev/tty
```

I found the compressed keyboard map (by Sarantos, dated 21-June-1997), including comments on its use, in

```
/usr/share/kbd/keymaps/i386/qwerty/gr.map.gz
```

The console font was stored in

```
/usr/share/kbd/consolefonts
```

where there are also several gr737 and gr928 fonts. Note that the ISO-8859-7 codes are translated to Unicode (UCS-2) for output (**man setfont**).

man bash and a quick search for input-meta, convert-meta and output-meta tells us that **input-meta** must be on for 8 bits, **convert-meta** must be off otherwise our 8th bit is translated into an escape character and **output-meta** must again be on for 8 bits.

Finally, the “Linux keyboard and console HOWTO – The console character sets” states that to **select the user-defined character sets** for display we must send the “ESC (K” and “ESC) K” sequences to the console (/dev/tty).

Following are the instructions provided in gr.map.gz by its author:

```
# Left-shift-alt is the greek-lock key (the same as in Win95).
# The new win95 keys (on win95 keyboard) can be used too
# (until a better use of them is found):
# Left and Right win95 keys act like shift-to-greek,
# and Right win95menu key acts like greek-lock.
#
# I was unable to define two dead keys in a row, so:
# For both accent and dialytika use Ctrl and the accent key.
#
# For bugs/remarks, send mail to sarantos@ics.forth.gr
# You can get updated versions from
# http://www.ics.forth.gr/~sarantos/gr.map
#
# This is an iso-8859-7 mapping. Use together with "setfont iso07.*"
# activated by sending Esc ( K to the console.
```

Now if you manage to reach a point where your console is unusable, try the following:

Press “Esc” followed by “c” or blindly type

```
echo “\033c” > /dev/tty/
```

Blindly type “**reset**” or “**setterm – reset**” or “**tput reset**” or “**tput init**” or “**tput clear**” at the console prompt. You may also have to type a line feed before and after these commands:

e.g. Press **Ctrl/J**, type **reset**, press **Ctrl/J** again.

Loop to initialize all terminals (for startup files, experts only):

```
for tty in /dev/tty[0-5]*;
        do echo -ne "\033(K\033)K" >$tty;
done
```

No polytonic keyboard

Sometimes one may wish to just enter a few characters, without having to grapple with keyboard maps and such. Enter the **character map** programs.

gucharmap by Noah Levitt and **kcharselect** by Reginald Stadlbauer will be of great help. The former includes Unicode character details, as well.

Word-processing and display applications

Just a few highlights – most contemporary GNU / Linux applications should be able to handle Unicode. If this is not the case, please report the problem(s) to their maintainers.

OpenOffice handles Hellenic accented characters well. Its search and replace function includes an option for using powerful **regular expressions** (**grep**-like syntax). It will produce **html** and **pdf** files using UTF-8 encoding.

Plain text editors: **xemacs** (21.5.x), **kate** (2.4.x - kate.kde.org), **kwrite** (4.4), **yudit**.

xpdf (3.00) will display UTF-8 and will allow you to enter Hellenic characters in the search box. Adobe **Acrobat 5.0.x** for Linux used to simply fold when it encountered a UTF-8 environment, otherwise it would show the text. No search with Hellenic, though. **Acrobat 7.0** for Linux works and allows for basic Hellenic characters in its search box. No extended characters (with breathings and accents).

Browsers (**konqueror**, **firefox**, **mozilla**) will happily display UTF-8 codes, as long as **charset** is defined and fonts match.

GNU Free Documentation License

Version 1.2, November 2002

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors

of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that

translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download

using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
D. Preserve all the copyright notices of the Document.
E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
H. Include an unaltered copy of this License.
I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties--for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions,

provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements."

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

How to Apply These Terms to Your New Programs

If you develop a new program, and you want it to be of the greatest possible use to

the public, the best way to achieve this is to make it free software which everyone can redistribute and change under these terms.

To do so, attach the following notices to the program. It is safest to attach them to the start of each source file to most effectively convey the exclusion of warranty; and each file should have at least the "copyright" line and a pointer to where the full notice is found.

```
    <one line to give the program's name and a brief idea of what it does.>
    Copyright (C) <year>  <name of author>

    This program is free software; you can redistribute it and/or modify
    it under the terms of the GNU General Public License as published by
    the Free Software Foundation; either version 2 of the License, or
    (at your option) any later version.

    This program is distributed in the hope that it will be useful,
    but WITHOUT ANY WARRANTY; without even the implied warranty of
    MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the
    GNU General Public License for more details.
    You should have received a copy of the GNU General Public License
    along with this program; if not, write to the Free Software
    Foundation, Inc., 51 Franklin St, Fifth Floor, Boston, MA  02110-1301
USA
```

Also add information on how to contact you by electronic and paper mail.

If the program is interactive, make it output a short notice like this when it starts in an interactive mode:

```
    Gnomovision version 69, Copyright (C) year name of author
    Gnomovision comes with ABSOLUTELY NO WARRANTY; for details type `show w'.
    This is free software, and you are welcome to redistribute it
    under certain conditions; type `show c' for details.
```

The hypothetical commands `show w' and `show c' should show the appropriate parts of the General Public License. Of course, the commands you use may be called something other than `show w' and `show c'; they could even be mouse-clicks or menu items--whatever suits your program.

You should also get your employer (if you work as a programmer) or your school, if any, to sign a "copyright disclaimer" for the program, if necessary. Here is a sample; alter the names:

```
  Yoyodyne, Inc., hereby disclaims all copyright interest in the program
  `Gnomovision' (which makes passes at compilers) written by James Hacker.
```

```
<signature of Ty Coon>, 1 April 1989
Ty Coon, President of Vice
```

This General Public License does not permit incorporating your program into proprietary programs. If your program is a subroutine library, you may consider it more useful to permit linking proprietary applications with the library. If this is what you want to do, use the GNU Library General Public License instead of this License.