New2Linux: Editing Text Files

September 2, 2021

John Nash
for
Virtual Meeting of Linux-Ottawa

Caution!

TEXT EDITING and WORD PROCESSING

- VERY different in the "product"

- TEXT – think of what a typewriter can produce
    - Fixed width of characters and fixed line spacing

- WP – think typesetting with fonts, glyphs, images, line spacing etc. -- WYSIWYG

- THIS TALK ABOUT TEXT EDITING

Why?

- Nearly all Linux/Unix and (GASP!) Windows systems use configuration files and scripts (small programs) that are plain text.
    - Perl, Python, ?ash (bash, csh, zsh, dash, ...), R, etc.
    - *.ini, *.cfg, etc
- Many important "documents" are plain text
    - (La)TeX
    - (?)Markdown
    - XML
    - HTML
- *ALL* users eventually have to edit a file such as these
- Editorstan: heavy use --> factions (religions?) and wars

VI: what newbies need to know

- Written in mid-1970s by Bill Joy because **ed** is too difficult even for immortals to use – somehow became entrenched

- Designed to play nicely with VERY slow 300 bps terminals and commands matched Lear Siegler keyboard

   http://www.softpanorama.org/Editors/Vimorama/history.shtml

- If you REALLY must learn about it:

   https://www.tutorialspoint.com/unix/unix-vi-editor.htm

- Otherwise remember

   ESC : q !

   so you can get out of VI (in order to run EMACS?).

What editor features are important?

- Relatively "obvious" operation

  - Depends hugely on previous experience of user

- Should be usable WITHOUT mouse/pointer, but for GUI editors, the pointer function should be operational

- Search and replace

  - With special characters? e.g., \t \n (gedit family)

- Undo/Redo (Ctrl-Z/Ctrl-Shift-Z)

- Offer line numbering and line wrap

- For modern editors, capable of working with different character encodings automatically

- Colour highlighting? -- only if you work with selected file types

Character encodings?

- The way computers translate strings of bits into alphabetic or other symbols

  - ASCII – a 7-bit encoding evolved from 1870s 5-bit Baudot telegraphic code

  - EBCDIC – an 8-bit encoding pushed by IBM

  - Unicode ==> UTF-8, UTF-16, UTF-32

    - Modern but messy set of standards

    - Defines a number for each character, but not its graphic form – LOTS of room for fights

    - Windows uses UTF-16; Linux / Web use UTF-8

Should I worry?

I do NOT try to get very involved in encodings, but

http://www.text-editor.org/encoding.htm

- Working ONLY in Linux in English? – No.

- Otherwise? -- Maybe.
  - You get files from Windows users
  - You have correspondents using non-English languages
  - You are doing a thesis in Linguistics

- DO want editor to be capable of dealing with encodings.

A nasty legacy – line endings

Normal text files end the lines as follows:

- Windows: CR+LF   (0D 0A hex, 13 10 dec)

- Linux: LF (0A hex, 10 dec)

- OS X: CR (0D hex, 13 dec)

"Most" modern text editors can cope automatically, BUT ...

- My Google Summer of Code student used Windows, and git kept "updating" after each of us viewed a file.

- Not a critical issue, but an annoyance

Column (block mode) editors?

- Possibly an esoteric topic

- For novices, message is that this is NOT a common capability, and rather tricky

- Very useful for some data management applications

- Some commercial offerings (UltraEdit)

- Linux: several possibilities. I'd choose geany?

  https://askubuntu.com/questions/6451/true-column-mode-block-selection-and-editing-text-editor-solution

What choice?

https://texteditors.org/cgi-bin/wiki.pl

- Most distros have a default choice, but you may not like it. e.g., I like Leafpad, but xed is Linux Mint default

- But even then, it may not be in repository

  - Find alternative repo, or snap, or build from source.

- For some tasks (e.g., editing tabs to spaces) more power needed, so I use gedit or a descendant (e.g. pluma, xed), but xed has lots of features I don't use, and does NOT apparently handle tab (\t) etc.

What about CLI mode?

- Character-based editors for use in terminal

- Experts!: vi vs. Emacs families (editor wars)

- For those who don't: pico, nano, joe

  - I like joe (as seeminigly does Linus Torvalds)

  - But I (we?) have a Turbo Pascal heritage – you like what you're used to

  - Generally do NOT use pointer but keystrokes to navigate the file.

Setting a default editor

- Some platform dependence

- GUI: Ubuntu/Mint use

  - file manager (Nautilus/Caja)

  - select file

  - Properties/Open With

- CLI: Debian family

  - sudo update-alternatives --config editor

Take away messages

- If you do a lot of text editing, it is worth choosing an editor that fits your needs and likes

- If using a lot of different machines and platforms:

  - Keep it simple. CLI use nano, GUI use gedit/pluma

  - Generally these or approximations available by default i.e., when you click on a text file

- BUT, if it is important to you, be willing to explore because there are plenty of fine choices