Foothill College
CS 30B Linux Shell Programming

Lecture Notes #3

This week we will look at some commonly used filters and regular expressions that will prove invaluable in your shell scripting tasks. Much of the power of Linux comes from utilities that operate on a file in a non-interactive mode to manipulate text. These utilities are commonly referred to as "*filters*" that read from standard input and write to standard output, allowing them to be used in pipes.

You can think of these utilities as *essential* tools to write useful and efficient scripts.

Three powerhouse utilities that will be discussed further down the road in the course include `grep`, `sed`, and `awk`. These programs are key Linux tools for pattern matching and manipulating text.

First, a solid foundation on the use of regular expressions and regular expression metacharacters will be discussed in order to put you in a position to exploit the power of the power players: `grep`, `sed` and `awk`.

In preview, the roles these different tools contribute to processing text include:

`regex` - is a language for describing patterns in strings
`grep`- filters its input against a pattern
`sed` - applies transformation rules to each line and
`awk` - manipulates an ad-hoc database stored as text (e.g. CSV files).
`tr` - translate characters – allows manipulation of the character of a file
`sort` – sort by field

It is best to begin with discussion in detail including examples along the way. Not to worry, we will begin small with specific topics on small data sets before unleashing the power tools.

Next week we will look at using command files, variables and built-in arithmetic in our scripts.