*UNIX – Shell Scripting*

Practice Handout – Regular Expression

Some Reading First :

Matching Words

If you want to match th, but only at the start of a word, you can precede it with \&<t;< tt="">, for example, \<th. Similarly, you can match it at the end of a word with \>, for example, th\>. Of course, you can match entire words, e.g., \<the\>.</t;<>

Alternation

Suppose, however, we wish to match the words {net, Net}. To allow multiple characters in one location, we can use [ ... ], adding any characters we are interested in between the brackets: [Nn]et.

Other examples of alternation are n[aeiou]t which matches {nat, net, nit, not, nut} or [Nn][aeiou]t which matches {nat, net, nit, not, nut, Nat, Net, Nit, Not, Nut}.

As you can see, you can very quickly describe, with a small number of characters, a large possible collection of strings.

To match [ and ], escape the character with a \.

1. There is a file hamlet.txt in your dropbox. Copy the content and create a file in your unix machine. Try this :

**egrep be hamlet.txt**

**egrep "[Bb]e" hamlet.txt**

**egrep "\<[Bb]e\>" hamlet.txt**

e**grep -n "\<[Bb]e\>" hamlet.txt**

1. Find all two-letter words ending with e by using
2. Find all matches of 'd\.’ and all matches for 'd.’
3. How would you match the following patterns?

hw1

hw2\*

/hw1

#include

1. Write down a single regular expression that will match any of the following

sn2a3

sn22a3

sn2b3

sn22b3