IT 240
Shell Scripting for Administrators

Chapter 9
Processing Text

Stan J. Senesy
IT Program/CCS
New Jersey Institute of Technology

Text Substitution

- Just as we used m// for pattern matching, we can use s/// for matching with substitution

  s/Barney/Fred

- If the string ‘Barney’ is found, it is replaced with ‘Fred’ and the return value is true (in a conditional test)

Global Replacement

- By adding /g to the end of a substitution, it becomes possible to replace all instances of a search text in a string:

```
$_ = “home, sweet home”
s/home/cave/g
```

Modifiers to s//

- The modifiers to m// work the same with substitution:
- To remove case sensitivity, use the /i qualifier
- To match any pattern over multiple lines, use the /s qualifier
- The /x modifier allows you to add arbitrary whitespace to a pattern; tabs or spaces inside a pattern are ignored

More Operators

- The binding operator ~= allows the substitution of the string on the right into the variable or string on the left
- You can modify the case of string characters as well:
  - The \U escape forces what follows to uppercase.
  - \E turns off case shifting
  - When lowercase \u and \e only affect the next character
  - \u and \L means all lower case but the first character

More Operators

- split is used to break a string up according to a pattern:

```
@fields = split /separator/, $string;
```

- Just as split breaks things up, join may be used to put them together:

```
my $result = join $glue, @pieces;
```

Advanced RegExp's

- Everything that we've looked at so far may be considered 'greedy' algorithms:
- In other words, they try to match as much as they can, and only reluctantly give something back if it's necessary to succeed
- Page 127 - backtracking!

Advanced RegExp

- For every greedy quantifier, a non-greedy version exists as well, defined by adding the ? to the quantifier
- When using a non-greedy quantifier, the match tries to use the minimum amount of characters to do the match, not the maximum