IT 240
Shell Scripting for Administrators

Chapter 7
Regular Expressions

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

Regular Expressions

- We’ve had an introduction to regex’s in the first part of the course
- To Perl, a regex is a pattern to be used principally for matching
- Patterns can match one or more strings
- Regex are not ‘globs’ (wildcard matches from shell scripts)

Simple Patterns

- Once we’ve defined a string, we can match a part or whole of it by putting the search pattern between the / and /

```
$_ = “yabba dabba doo”;
if (/abba/) {
    print “abba matched!\n”;
}
```

Meta-characters

- Perl provides it’s own wildcards
- The . matches any single character
- The \ is also a metacharacter and may be used to ‘normalize’ any other meta-character (for instance, when searching for a .)

Meta-characters

- The * means to match the preceding item zero or more times
- The + sign may be used to match the preceding item one or more times
- The ? indicates that the preceding character is optional and that it may or may not occur

Meta-characters

- Parenthesis are used for grouping and are also considered meta-characters
- The | means that either the expression on the left or right side of the bar may match

Character Classes

- A class is a listing of possible characters contained inside the [ ]
- It matches one of the characters, but it may be any of the characters in the list
- For example:

  [a-zA-z]

  [0-9]

  [abcdefg]

Shortcuts

- Specifying common classes can get a bit repetitious, so Perl provides shortcuts:

  \w - [A-Za-z0-9_]

  \s - whitespace

  \d - [0-9]

- Shortcuts may be negated by using it's uppercase counterpart