ULI101: INTRODUCTION TO UNIX / LINUX AND THE INTERNET

WEEK 9: LESSON 2

EXTENDED REGULAR EXPRESSIONS
LINUX COMMANDS THAT USE REGULAR EXPRESSIONS

LESSON 2 TOPICS

Extended Regular Expressions

- Definition / Purpose

- Extended Regular Expressions Symbols

- Instructor Demonstration

Other Linux Commands That Use Regular Expressions

- `man` , `more` , `less` , `vi` , `sed` , `awk`

Perform Week 9 Tutorial

- Investigation 2

- Review Questions (<u>**Extended**</u> **Regular Expressions**, Parts **A** and **B**)

EXTENDED REGULAR EXPRESSIONS

Extended Regular Expressions

Extended Regular Expressions consist of additional special characters that "**extend**" the capability of regular expressions.

We will discuss three types of extended regular expressions:

     Repetition: `{min, max}` , `?` , `+`

     Grouping: `(   )`

     `Or` **Condition:** `|`

EXTENDED REGULAR EXPRESSIONS

AAAA

Repetition

The extended regular expression symbol consists of the
minimum and/or **maximum** number of repetitions contained within braces `{  }`

Usage:

`{min,max}`

Examples:

`a{2,5}`   **2 to 5** occurrences of the character **a**

`[0-9]{1,}`   **1 or more** occurrences of a **number**   no upper limit
`[0-9]+`   (shortcut method)

`[a-z]{0,1}`   **zero or 1** occurrence of a **lowercase letter**
`[a-z]?`   (shortcut method)

EXTENDED REGULAR EXPRESSIONS

AAAA

Repetition Extended Regular Expression Example

If you issue the **grep** command <u>without</u> options with
extended regular expressions, the command **will NOT work**.

When using the grep command with extended regular
expressions you *need to use* **egrep** *or* **grep -E**

Examples: greedy match: try match as much as possible

```
egrep "^[0-9]{1,}$" data.txt
egrep "^[+-]{0,1}[0-9]{1,}$" data.txt
egrep "^[0-9]{1,}[.]{0,1}[0-9]{0,}$" data.txt

grep -E  "^[0-9]{1,}$" data.txt
grep -E "^[+-]{0,1}[0-9]{1,}$" data.txt
grep -E "^[0-9]{1,}[.]{0,1}[0-9]{0,}$" data.txt
            [a]{1,} equal to 'aa*'
```

```
cat data.txt
123
+45
+++37
-67.89
--57.6
-78...4
12.6
+26.887


egrep "^[0-9]{1,}$" data.txt
123


egrep "^[+-]{0,1}[0-9]{1,}$" data.txt
123
+45


egrep "^[0-9]{1,}[.]{0,1}[0-9]{0,}$" data.txt
123
12.6
```

EXTENDED REGULAR EXPRESSIONS

(pattern)

Grouping

If you want to search for repetition for a **group** of **characters**, a **word**, or a **phase**, you can enclose them within brackets **( )**

Examples:

two "the", nothing in between them

```
egrep "(the )(2,}" data.txt
egrep "(lazy fox )(2,3}" data.txt
```

```
cat data.txt
The lazy fox jumped over dog
Time to go to the the store
I like to go to the movies
I act like a lazy fox lazy fox lazy fox
Don't be a lazy fox

egrep "(the ){2,}" data.txt
Time to go to the the store

egrep "(lazy fox ){2,3}" data.txt
I act like a lazy fox lazy fox lazy fox
```

lazy fox 5 times in a row still matches
as there is no anchor

REGULAR EXPRESSIONS

Or **Condition**

(this |that)

The **|** symbol is used as the "or" symbol to provide **alternatives** within a **group**.

Examples:

```
egrep "(this | that ){1,}" data.txt
egrep "(a|b|c){3,}" data.txt
        similar to [a,b,c]
```

```
cat data.txt
I know this is the day
Because that is correct
We don't know that it is sunny
I know how to cccamp
I waaaaant a tissue
Can a bbborrow a cup of sugar?

egrep "(this | that ){1,}" data.txt
I know this is the day
Because that is correct
We don't know that it is sunny

egrep "(a|b|c){3,}" data.txt
I know how to cccamp
I waaaaant a tissue
Can a bbborrow a cup of sugar?
```

REGULAR EXPRESSIONS

Instructor Demonstration

Your instructor will demonstrate examples of using
Extended Regular expressions with the **egrep** command.

    egrep this|that cars (doesn't have that command)
    egrep 'ford|chevy' cars
    egrep '(ford|chevy)' cars

REGULAR EXPRESSIONS

Other Linux Commands that Use Extended Regular Expressions

There are other Linux commands / utilities in addition to *grep* or *egrep* that use regular expressions.

You have already used a few of these commands like:
man, **more**, **less**, and **vi**.

Other commands like **sed** and **awk** will be taught in a <u>future</u> lesson.

REGULAR EXPRESSIONS

**Other Linux Commands that Use
Extended Regular Expressions**

man , **more** , **less**

When searching for patterns using the **man**, **more**, or **less**
commands**,** you specify a regular expression with a **forward slash /**

*Example with **man ls** command:*

/classify

```
-F, --classify
        append indicator (one of */=>@|) to entries

--file-type
        likewise, except do not append '*'

--format=WORD
        across -x, commas -m, horizontal -x, long -l,

--full-time
        like -l --time-style=full-iso
```

REGULAR EXPRESSIONS

**Other Linux Commands that Use
Extended Regular Expressions**

vi

The **vi** text editor use regular expressions to search
and manipulate (edit) text within a text document.

Examples:

/pattern   – search for pattern in text file
:%s/uli101/ULI101/g   – search and replace text globally (all lines)

```
I am taking the course ULI101
There are a lot of commands taught in ULI101
I am over half way in the ULI101 course
I like Linux


~
~
~
~
```

REGULAR EXPRESSIONS

Using Regular Expressions with Linux Commands other than grep

`awk` , `sed`

The **awk** and **sed** Linux utilities are used to **manipulate** text, from files or via Linux pipeline commands.

You will learn how to use these commands in a **later** lesson.

REGULAR EXPRESSIONS

Instructor Demonstration

Your instructor will demonstrate examples of using **Extended Regular Expressions** with the **man**, **more**, **less** and **vi** utilities.

HOMEWORK

Getting Practice

Perform **Week 9 Tutorial:**
(Due: Friday Week 10 @ midnight for a 2% grade):

- INVESTIGATION 2: EXTENDED REGULAR EXPRESSIONS

- INVESTIGATION 3: OTHER COMMANDS THAT USE REGULAR EXPRESSIONS

- LINUX PRACTICE QUESTIONS

    (**Extended Regular Expressions**, Parts **A** and **B**)