

## Homework 3: CS114 Unix Tools, Fall 2007

Due electronically (via CMS) on Wednesday 17 October 2007, 11:59PM

### 1 Softlinks

The file `/home/ejb34/HW3/start` is a softlink. What is the full absolute path of the file that `start` points to? That file is a softlink also; what is the full absolute path that it points to? Continue this chain until you reach a real directory.

Turn in a file `softlinks.txt` containing, one per line, in order, the full paths of each file in this chain.

### 2 Mountpoints

Recall that `mount` with no arguments prints all disks on the current machine and where in the filesystem they are mounted. For each of the following files, indicate what the name of the device is (e.g. `/dev/sda25`) that the file is actually on, and the path relative to the mountpoint. For example, if `/dev/sda2` is mounted on `/boot`, then for the path `/boot/foo/bar` you would write `"/dev/sda2 foo/bar"`.

```
/usr/bin/yes
/etc/passwd
/var/mail/root
```

Turn in a file `mountpoints.txt` containing, one per line, the mountpoints and paths for this section.

### 3 Pipelines

Write a one-line command which prints out all users on empire who have a user id (UID) less than 5. The information you need is in the file `/etc/passwd`, and you can find out about the format of this file by doing `man 5 passwd` (the 5 means you're looking in "section 5" of the manual). Your command should not rely on the value of the current directory, and should be a single pipeline. Helpful commands: `tr`, `awk`. The result should be `root`, `bin`, `daemon`, `adm`, and `lp` (that is, it should print those five usernames, one per line, to standard output, in some order).

Turn in a file `pipeline.txt` containing the command you write here. Make sure to test the command!

### 4 Regular Expressions

A web url is a string of the form `http://domainname/some/complicated/path`.

For example, the path to this file is `http://www.cs.cornell.edu/courses/cs114/2007fa/HW3.pdf`,

in which the domain is `www.cs.cornell.edu` and the path is `courses/cs114/2007fa/HW3.pdf`. Your goal in this problem is to write a `sed` command to extract domain names.

Specifically, write a pipeline which reads from a file `urls` and prints to standard output every unique domain name that appears in a url in that file. You may assume that there is only one url per line (if this problem is easy for you, consider how to handle multiple urls per line). Each domain name should appear only once in your output.

For example, if `urls` contains this text

```
Eric is teaching http://www.cs.cornell.edu/courses/cs114 but
last semester he TA'd http://www.cs.cornell.edu/courses/cs312
A very useful resource is http://www.google.com
Good luck!
```

your pipeline should print out

```
www.cs.cornell.edu
www.google.com
```

Useful commands for this problem: `sed`, `grep`, `sort -u`.

Turn in a file `regex.txt` containing your solutions to this problem.