

Introduction to Unix

Part II: Special Topics

Frederick J Tan
Bioinformatics Research Faculty
Carnegie Institution of Washington, Department of Embryology

16 April 2013

A Three Hour Tour

Part I: The Essentials

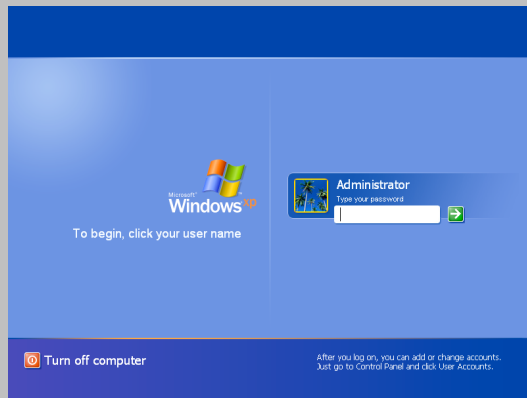
client-server model, command-line interface, navigation,
viewing files, searching, finding help

Part II: Special Topics

system administration, installing programs, multi-tasking
account settings, programming

How All the Pieces Fit Together

Your Physical Laptop



Virtual Server



```
Ubuntu 12.10 ubuntu tty1
```

```
ubuntu login: _
```

Local Commands



```
smith@laptop$ scp  
fedora:file.txt Desktop
```

Remote Control



```
smith@laptop$ ssh ubuntu
```



```
smith@ubuntu$ []
```

Department Server



```
Fedora release 9 tty1
```

```
fedora login: _
```

Catching Up to the Same Page

restrict to
start of line
↓

pipe
↓

split at
'A's
↙ ↘

print
cols 1-4
↓

continue on
next line
↙

```
$ grep ^GATTACA reads.fastq | cut -dA -f1,2,3,4 \
--output-delim=' ' | sort -k4 | uniq -c > gattaca-motif.txt
```

↑
tell cut to separate
cols with a space

↖ ↗
sort on bases
after GATTACA

↖ ↗
count
unique

↖
save output
to file

```
$ scp hg19-chr22-UCSC-knownGene.bed workshop@192.168.56.101:
```

Part II - Special Topics

System Administration

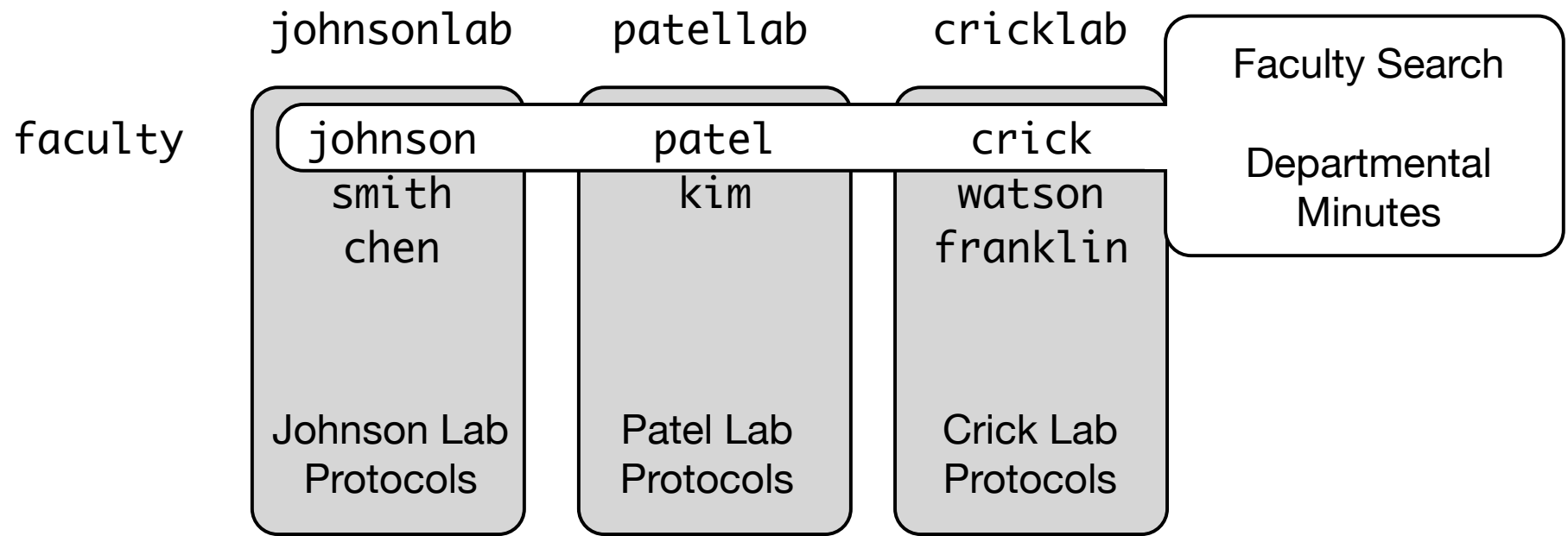
Installing Programs

Multi-Tasking

Account Settings

Programming

Security - Roles: Users and Groups



USER
johnson
smith
chen
patel
...

GROUP
johnsonlab, faculty
johnsonlab
johnsonlab
patellab, faculty
...

Security - Permissions:

user/group/other and read/write/execute

```
/home/workshop/bodymap$ ls -l
total 272M
drwxrwxr-x 2 workshop workshop 4.0K Apr  5 17:50 data
-rw-rw-r-- 1 workshop workshop 368 Apr 15 18:41 gattaca-motif.txt
-rw-rw-r-- 1 workshop workshop 272M Apr 15 18:40 reads.fastq
```

↑ ↑ ↑
permissions user group

d	directory
r	read
w	write
x	execute

drwxrwxrwx
 ↑ ↑ ↑
user group other

Task 1: Change File Permissions

```
$ chmod g-w,o-r,o+x gattaca-motif.txt
$ ls -l
-rw-r---x 1 workshop workshop 368 Apr 15 18:41 gattaca-motif.txt
```

```
$ groups
workshop adm cdrom sudo dip plugdev lpadmin sambashare
```

```
$ chgrp adm gattaca-motif.txt
$ ls -l
-rw-r----- 1 workshop adm 368 Apr 15 18:41 gattaca-motif.txt
```

```
$ chown games gattaca-motif.txt
chown: changing ownership `gattaca-motif.txt':Operation not permitted
```

Task 2: Respect the Power of `sudo`

```
$ sudo chown games gattaca-motif.txt
[sudo] password for workshop:
$ ls -l
-rw-r----- 1 games      adm          368 Apr 15 18:41 gattaca-motif.txt
```

```
$ touch gattaca-motif.txt
touch: cannot touch `gattaca-motif.txt': Permission denied
```

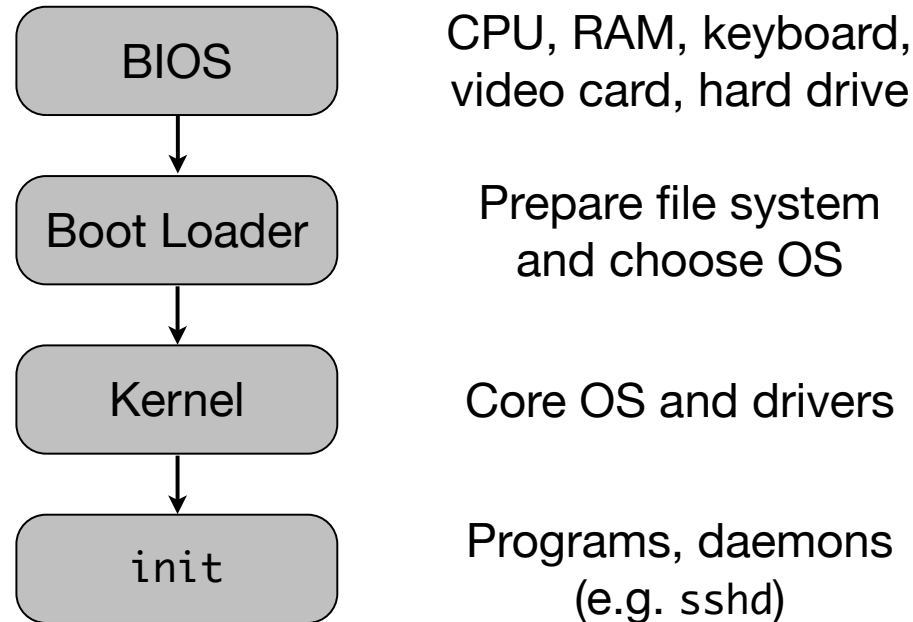
```
$ ls -l /etc/shadow
-rw-r----- 1 root shadow 717 Apr  5 16:51 /etc/shadow
```

```
$ cat /etc/shadow
cat: /etc/shadow: Permission denied
```

```
$ sudo cat /etc/shadow
...
workshop:$6$Kkpq51me$M3fE8nj1GAXU6WrsFp0mIT8gNRRzIgptc11sQAWlTk37g
```

```
$ sudo shutdown -h now
```

Booting Up a Unix System



```
$ ping 192.168.56.101
```

Part II - Special Topics

System Administration

Installing Programs

Multi-Tasking

Account Settings

Programming

Source Code vs Machine Code

Easy for a human to read

```
#include <stdio.h>

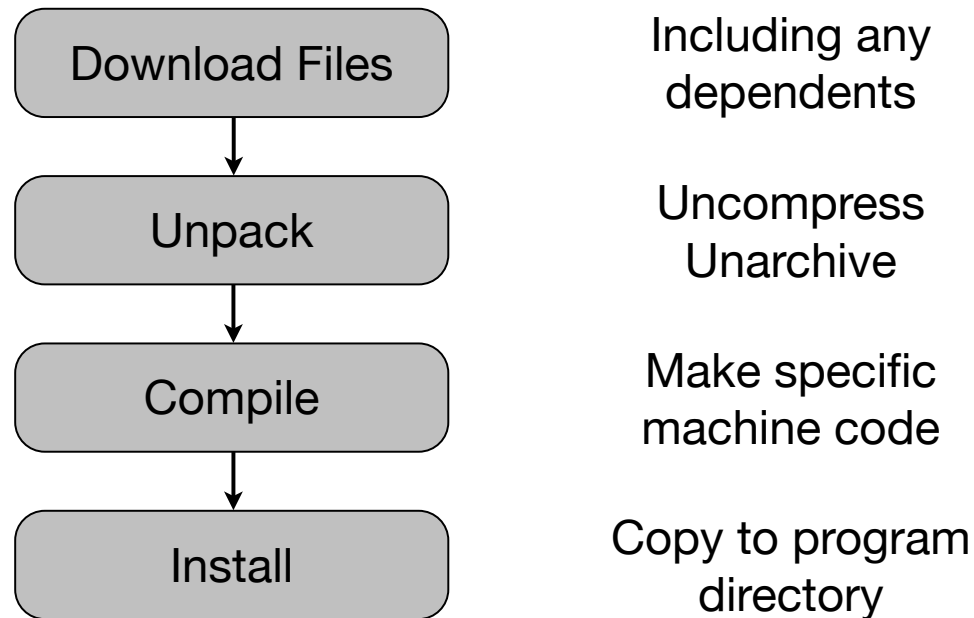
main()
{
    printf("Hello, world!");
}
```

compiler/
interpreter

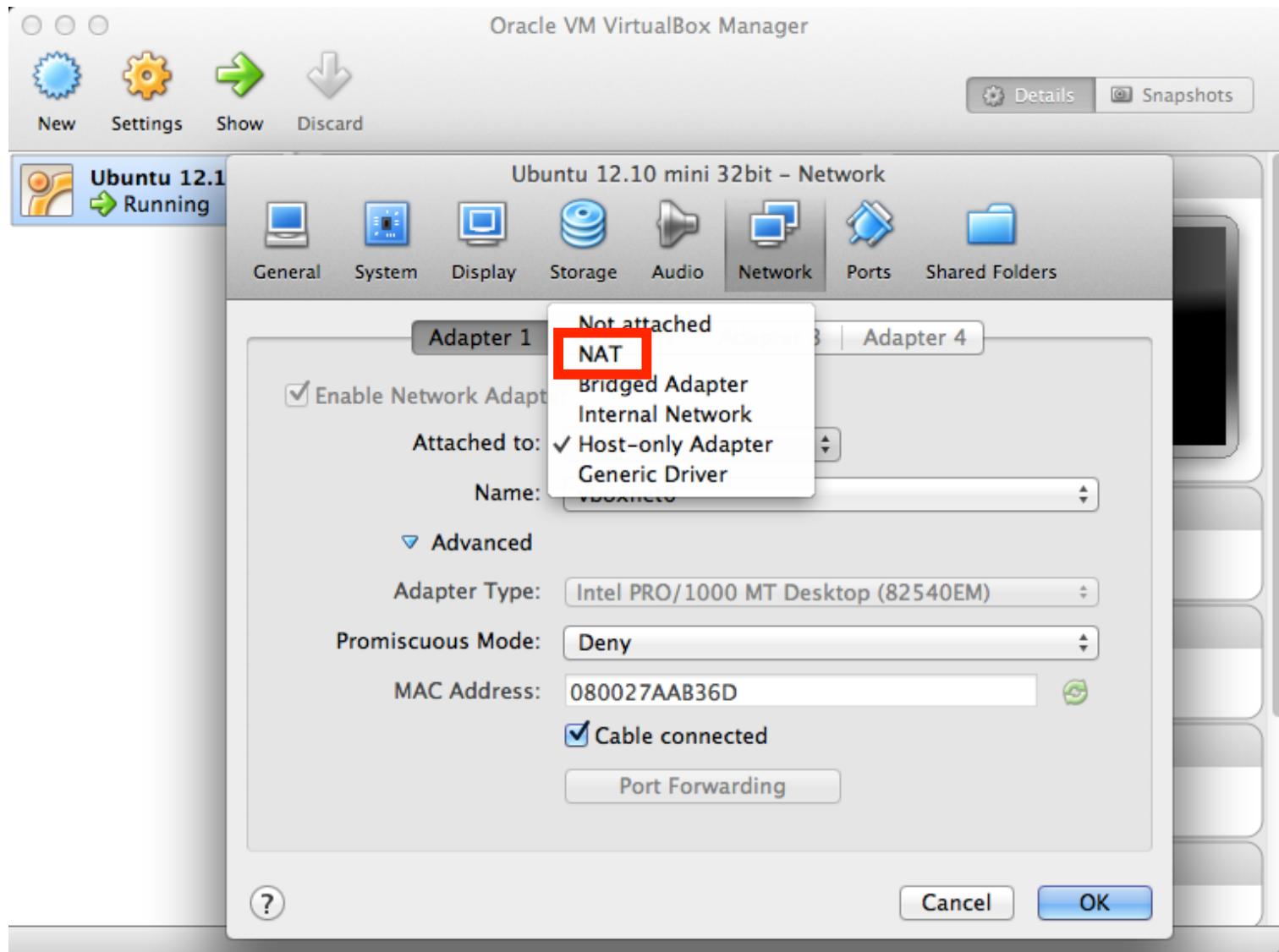
Easy for a computer to read

00004a0:	f3c3	8b1c	24c3	6690	5383	ec08	e800	0000	\$.f.S.....
00004b0:	005b	81c3	4f1b	0000	83c4	085b	c300	0000	.[..0.....[....
00004c0:	0300	0000	0100	0200	4865	6c6c	6f2c	2077	Hello, w
00004d0:	6f72	6c64	2100	0000	011b	033b	3000	0000	orld!.....;0...
00004e0:	0500	0000	08fe	ffff	4c00	0000	34ff	ffff	L...4...
00004f0:	7000	0000	58ff	ffff	9000	0000	c8ff	ffff	p...X.....

Common Steps to Install a Program



Do Try This (Later) at Home



Automagically Install System-Wide Programs with `apt-get`

```
$ apt-cache search ^hello
hello - The classic greeting, and a good example
hello-debhelper - The class greeting, and a good example
```

```
$ apt-get install hello
E: Could not open lock file /var/lib/dpkg/lock
E: Unable to lock the administration directory, are you root?
```

```
$ sudo apt-get install hello
Reading package lists... Done
Building dependency tree
The following NEW packages will be installed:
  hello
Get:1 http://us.archive.ubuntu.com/ubuntu/ quantal/main hello i386
Unpacking hello (from ../archives/hello_2.8-2_i386.deb) ...
Processing triggers for install-info ...
Setting up hello (2.8-2) ...
```

Task 3: Manually Compile and Install samtools

```
$ cd ~/src  
$ ls  
bowtie-0.12.9          samtools-0.1.19  
bowtie-0.12.9-linux-i386.zip  samtools-0.1.19.tar.bz2
```

```
$ rm -r s<TAB>  
$ rm ~/bin/s<TAB>
```

```
$ bunzip2 s<TAB>
```

```
$ tar xf s<TAB>
```

```
$ cd s<TAB>
```

```
$ less INSTALL
```

```
$ make
```

```
$ cp samtools ???
```

Unpack

Compile

Install

Understanding \$PATH Resolution

```
$ samtools
The program 'samtools' is currently not installed.
```

```
$ ./samtools

Program: samtools (Tools for alignments in the SAM format)
Version: 0.1.19-44428cd

Usage:  samtools <command> [options]
```

```
$ groups
workshop adm cdrom sudo dip plugdev lpadmin sambashare
```

```
$ whereis groups
groups: /usr/bin/groups /usr/share/man/man1/groups.1.gz
```

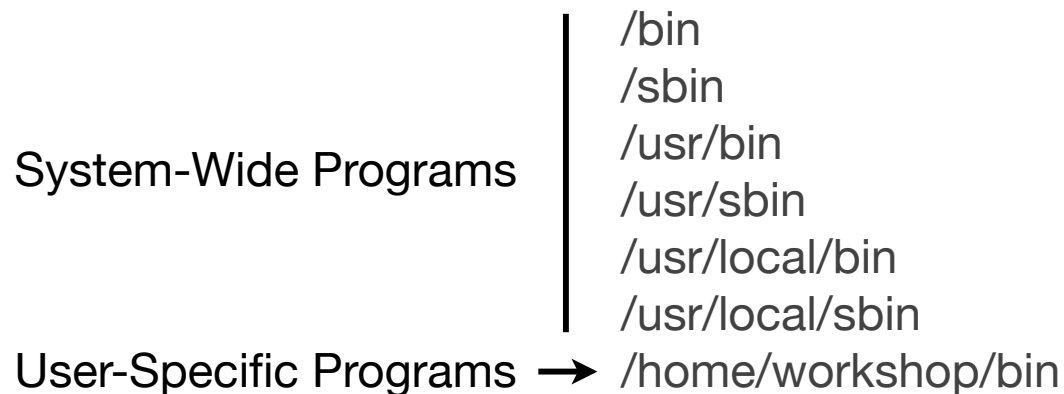
```
$ echo $PATH
/home/workshop/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin
:/usr/bin:/sbin:/bin:/usr/games:/usr/local/games
```

Task 4: Install a User-Specific Program

```
$ cp samtools /usr/bin
cp: cannot create regular file `/usr/bin/samtools': Permission denied
```

```
$ ls -l /usr
drwxr-xr-x  2 root root  20K Apr 15 20:33 bin
```

```
$ echo $PATH
/home/workshop/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin
:/usr/bin:/sbin:/bin:/usr/games:/usr/local/games
```



```
$ cp samtools ~/bin
```

Part II - Special Topics

System Administration

Installing Programs

Multi-Tasking

Account Settings

Programming

Task 5: Run a Job in the Background

```
$ cd ~/bodymap
$ bowtie ~/genomes/human/hg19-chr22 reads.fastq > reads.txt
<CTRL>-C
```

```
$ <UP><SPACE>&<ENTER>
```

```
[1] 1885
```

```
$ top
```

```
$ jobs
```

```
$ ps raux | grep workshop
```

workshop	1885	99.0	7.4	107472	38120	pts/0	R	04:55	5:06	bowtie
workshop	1962	0.0	0.2	5204	1196	pts/0	R+	05:00	0:00	ps raux

```
$ kill 1885
```

```
$ jobs
```

```
[1]+  Terminated                  bowtie ~/genomes/human/hg19-chr22 re
```

Task 6: **Keep a Jobs Running** with `nohup`

```
$ shopt -s huponexit  
$ bowtie ~/genomes/human/hg19-chr22 reads.fastq > reads.txt &  
$ ps raux | grep workshop  
$ exit
```

```
$ ssh workshop@192.168.56.101  
$ ps raux | grep workshop
```

```
$ nohup bowtie ~/genomes/human/hg19-chr22 reads.fastq > reads.txt &
```

Part II - Special Topics

System Administration

Installing Programs

Multi-Tasking

Account Settings

Programming

Password, Accounts, and Shells

```
$ passwd
Changing password for workshop.
(current) UNIX password:
Enter new UNIX password:
Retype new UNIX password:
passwd: password updated successfully
```

```
$ sudo adduser smith
Adding user `smith' ...
Adding new group `smith' (1001) ...
Adding new user `smith' (1001) with group `smith' ...
Creating home directory `/home/smith' ...
```

~~Terminal.app, ssh, PuTTY, "DOS Window"~~
bash, sh, ksh, csh, tcsh, zsh, cmd.exe

```
$ chsh
Password:
Changing the login shell for workshop
Enter the new value, or press ENTER for the default
Login Shell [/bin/bash]:
```

Task 7: Set a bash Environment Variable

```
$ echo $PATH  
/home/workshop/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin  
:/usr/bin:/sbin:/bin:/usr/games:/usr/local/games
```

```
$ echo $path
```

```
$ foo=bar  
$ echo $foo  
bar
```

```
$ printenv  
TERM=xterm-256color  
SHELL=/bin/bash  
SSH_CLIENT=192.168.56.1 62302 22  
SSH_TTY=/dev/pts/0  
USER=workshop  
...
```

Task 8: Examine bash and System Configuration Files

```
$ cd
$ ls -a
.          .bash_logout  bodymap    HelloWorld.tar.gz      src
..         .bashrc       .cache     hg19-chr22-UCSC-knownGene.bed
.bash_history bin          genomes     .profile
```

```
$ less .bashrc
/PS1
/ls=
```

```
$ ls /etc/default
```

```
$ less /etc/default/grub
```

Part II - Special Topics

System Administration

Installing Programs

Multi-Tasking

Account Settings

Programming

Types of Programming Projects

Run a Series of Programs	shell script
Convert a .gff File to a .bed File	Perl, Python
Create a Better Motif Finder	C, C++, Java

Text Files vs Non-Text Files

```
$ gunzip H<TAB>
```

```
$ tar xf H<TAB>
```

```
$ cat H<TAB>.txt
```

```
Hello, world!
```

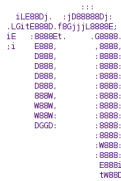
```
$ cat H<TAB>.rtf
```

```
{\rtf1\ansi\ansicpg1252\cocoartf1038\cocoasubrtf360
{\fonttbl{\f0\fswiss\fcharset0 Helvetica;}
{\colortbl;\red255\green255\blue255;}
\margl1440\margr1440\vieww9000\viewh8400\viewkind0
\pard\tx720\tx1440\tx2160\tx2880\tx3600\tx4320\tx5040\tx5760\tx64
80\tx7200\tx7920\tx8640\ql\qnatural\pardirnatural
\f0\fs24 \cf0 Hello, world!}
```

```
$ cat H<TAB>.doc
```

```
???>??%'????$????????????????????????????????????????
???????????????????????????????????????? bjbj?? $v????????qqqq
```

Most Text Editors Provide Syntax Highlighting



TextWrangler, UltraEdit, Emacs, Vim, nano

```
$ nano HelloWorld.c
```

```
tan — workshop@ubuntu: ~ — ssh — 80x24
GNU nano 2.2.6 File: HelloWorld.c Modified

/* Hello World program */

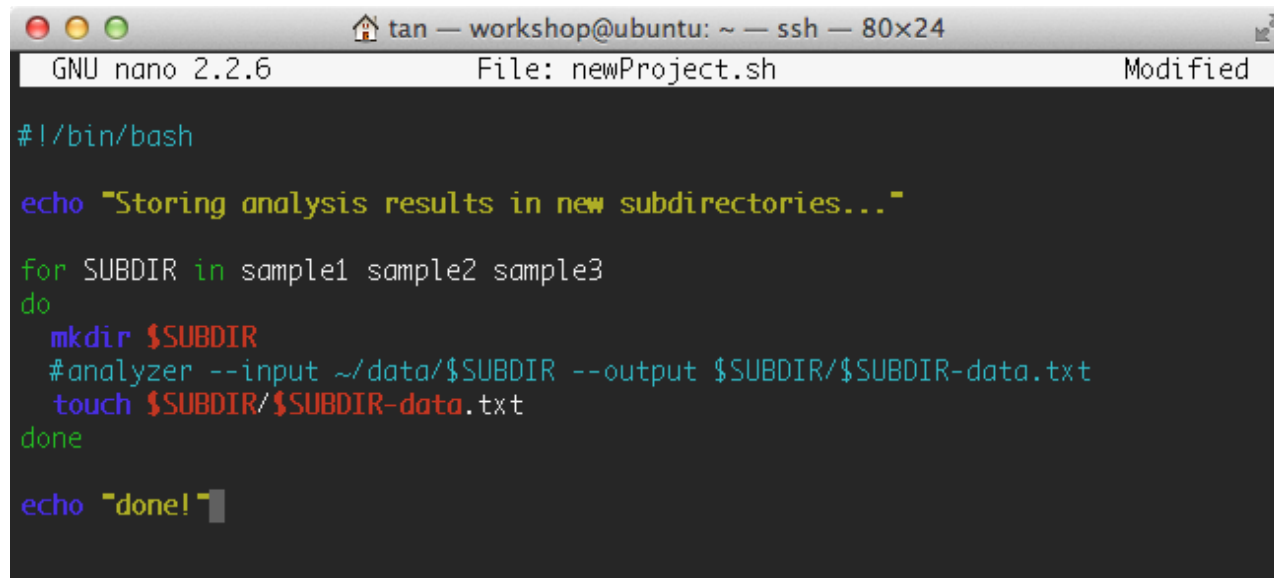
#include <stdio.h>

main()
{
    print("Hello, world!");
}
```

^G Get Help ^O WriteOut ^R Read File ^Y Prev Page ^K Cut Text ^C Cur Pos
^X Exit ^J Justify ^W Where Is ^V Next Page ^U UnCut Text ^T To Spell

Task 9: Make a Simple Shell Script

```
$ nano newProject.sh
```



```
tan — workshop@ubuntu: ~ — ssh — 80x24
GNU nano 2.2.6 File: newProject.sh Modified
#!/bin/bash
echo "Storing analysis results in new subdirectories..."
for SUBDIR in sample1 sample2 sample3
do
    mkdir $SUBDIR
    #analyzer --input ~/data/$SUBDIR --output $SUBDIR/$SUBDIR-data.txt
    touch $SUBDIR/$SUBDIR-data.txt
done
echo "done!"
```

Task 10: Run Your newProject.sh Script

Task 11: Make a `grep` Substitute in Perl

GOAL: find genes on + strand

SOLUTION A: `grep`

```
$ grep + hg19-chr22-UCSC-knownGene.bed | less -S
```

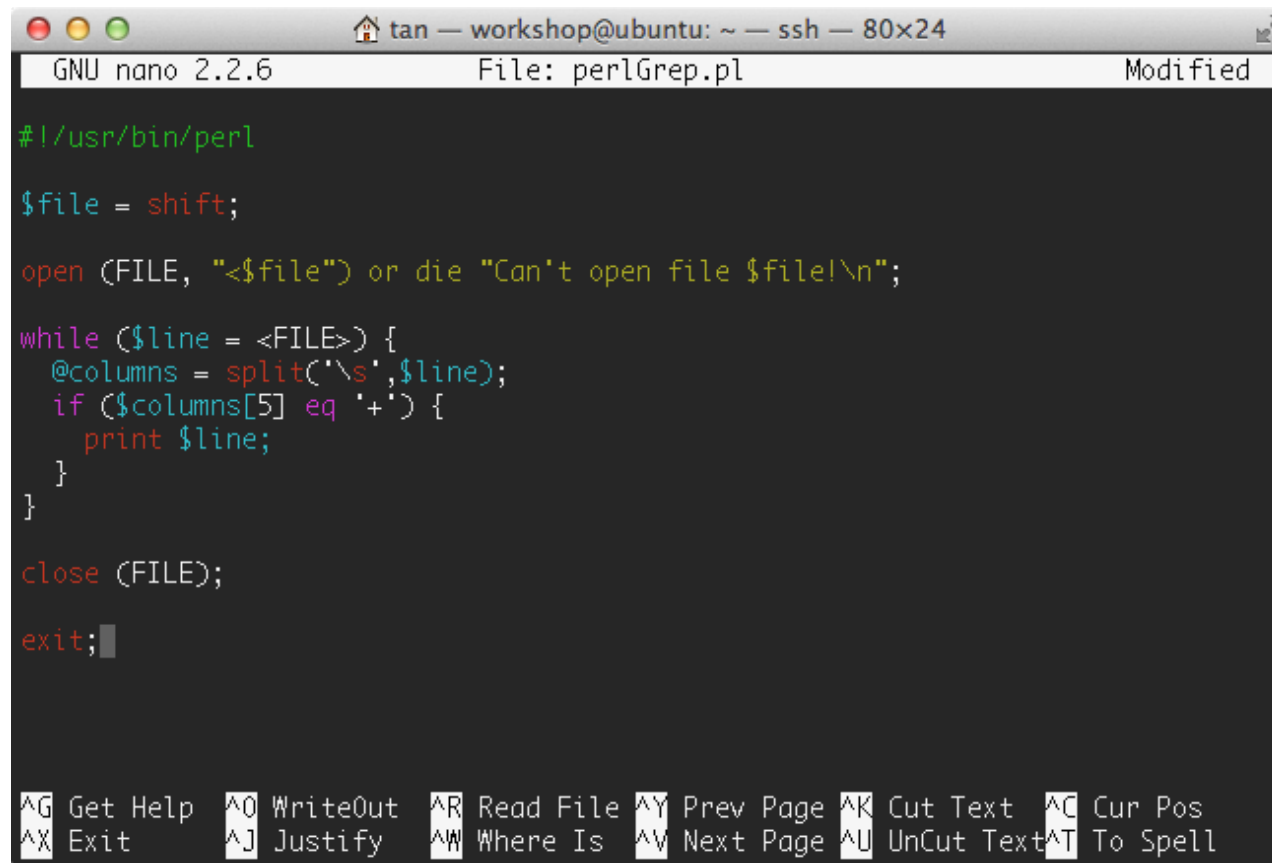
SOLUTION B: Perl script

PSEUDOCODE:

1. Obtain name of file to interrogate
2. Prepare file ... locate, open
 - > if can't find/open then report error!
3. Read a line from the file
4. Test if line contains a + symbol in column 6
 - > if yes, print entire line
5. If not at the last line, go to Step 3
6. Properly store file ... close
7. Report that everything proceeded as normal!

Task 11: Make a `grep` Substitute in Perl

```
$ nano perlGrep.pl
```



```
tan — workshop@ubuntu: ~ — ssh — 80x24
GNU nano 2.2.6 File: perlGrep.pl Modified

#!/usr/bin/perl

$file = shift;

open (FILE, "<$file") or die "Can't open file $file!\n";

while ($line = <FILE>) {
    @columns = split('\s',$line);
    if ($columns[5] eq '+') {
        print $line;
    }
}

close (FILE);

exit;
```

^G Get Help ^O WriteOut ^R Read File ^Y Prev Page ^K Cut Text ^C Cur Pos
^X Exit ^J Justify ^W Where Is ^V Next Page ^U UnCut Text ^T To Spell

```
$ ./perlGrep.pl hg19-chr22-UCSC-knownGene.bed | less -S
```

Task 12: Quick “One-Liners” in AWK

```
$ awk '{if ($6 == "+") print}' hg19-chr22-UCSC-knownGene.bed
```

PSEUDOCODE:

1. Read a line from stdin
2. Split line into columns, as demarcated by whitespace (tabs, space, etc.)
3. Test if column 6 is a + symbol
 - > if yes, print entire line
4. If not at the last line, go to Step 1
5. Report that everything proceeded as normal!

```
$ awk '{if ($6 == "+") print ($3-$2)}' hg19-chr22-UCSC-knownGene.bed
```

```
$ awk '{if (($6 == "+") && (($3-$2) > 10000)) print}' \
hg19-chr22-UCSC-knownGene.bed
```

A Three Hour Tour

Part I: The Essentials

client-server model, command-line interface, navigation,
viewing files, searching, finding help

Part II: Special Topics

system administration, installing programs, multi-tasking
account settings, programming

Useful Websites

Unix commands	linux.about.com/library/cmd/blcmdl1_export.htm
Unix utilities	en.wikipedia.org/wiki/List_of_Unix_utilities
Unix system directories	en.wikipedia.org/wiki/Unix_directory_structure
Unix habits	ibm.com/developerworks/aix/library/au-badunixhabits.html
grep	robelle.com/smugbook/regexpr.html
Unix questions	stackoverflow.com
Sys admin questions	serverfault.com
Bioinformatics questions	biostars.org
NexGenSeq questions	seqanswers.com
Workshop notes	emb.carnegiescience.edu/labs/frederick-tan