



# Before You Begin

Introduction to Unix for HPC

Raijin version

# Agenda – Day 1 – Shell Primer

9:45	Welcome
10:00 – 11:15	Session 1 (Intro & Unit 1)
11:15 – 11:30	Short break
11:30 – 12:30	Session 2 (Unit 2+3)
12:30 – 1:30	Lunch
13:30– 14:45	Session 3 (Unit 4)
14:45 – 15:00	Short Break
15:00 – 16:15	Session 4 (Unit 5)

# Agenda – Day 2 – Data & HPC

9:45 – 11:15	Session 1 (Data – Units 1+2+3)
11:15 – 11:30	Short break
11:30 – 12:30	Session 2 (Data – Units 4+5)
12:30 – 1:30	Lunch
13:30 – 15:30	Session 3 (HPC)

# General Introduction

- Intersect <http://www.intersect.org.au/>
  - Who we are?
  - Your Trainer
- Your University IT Contacts
- General Housekeeping
  - Toilets
  - Coffee & Water Facilities
  - Emergency Exits

# Course Information

- Who is the course intended for?
- What will be covered during the 2 days?
  - More details in the course outline at <http://www.intersect.org.au/course-resources>
- What level will you be at upon completion of the course?
- **Remember – It's YOUR course, so ask questions!**

# Download training materials

Training Material available from:

<http://www.intersect.org.au/course-resources>

- Download the Slides and Exercises and save to a local directory, e.g. C:\HPC:
  - **Course Outlines:** Shell Primer, Data, HPC
  - **Slides:** Shell Primer, Data, HPC Slides
  - **Exercises**

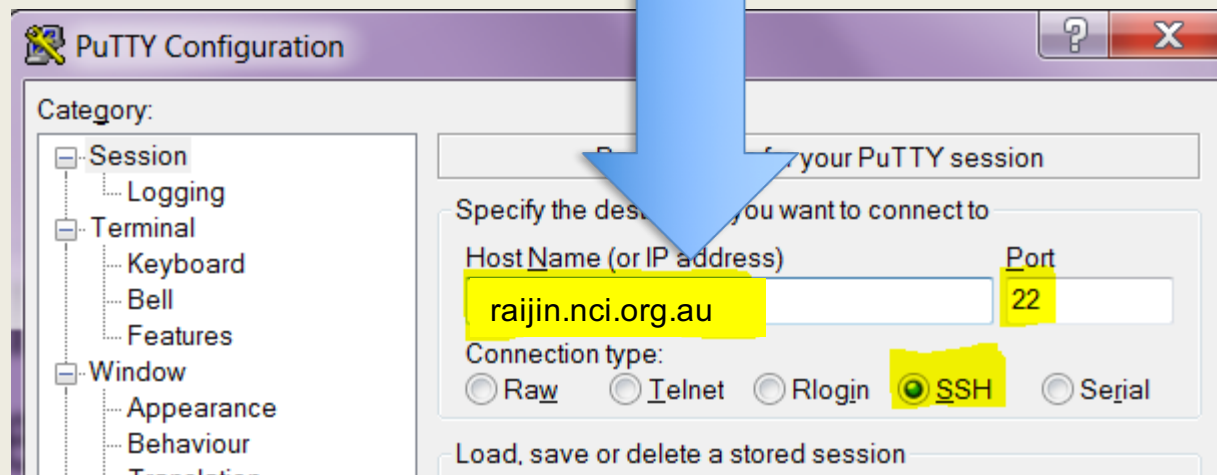
# Install software

Use your web browser to download Putty and PuttySCP from: <http://www.putty.org>

- Click on the “Download Putty” link and download:
  - putty.exe (a Telnet and SSH client)
  - pscp.exe (an SCP client, i.e. command-line secure file copy)
- Double click to install on your PC.

# Logging On

- Start the putty telnet/ssh client by double clicking on putty.exe and connect to the HPC Machine
  - Host: `raijin.nci.org.au`
  - Connection Type: `ssh`
  - Port: `22`





# Log On to Raijin

- Log into Raijin using the Test Account allocated to you, e.g.
  - **Account Name:** iaa444 – ibx444  
– iaa, iab, iac... iax, iba, ibb, ibc ...ibx
  - **Password:** provided by your instructor



# Shell Primer

Introduction to Unix for HPC

# Unit 1: Run Spot, Run

## Goals:

- Can run commands, including when there are spaces in arguments
- Can use 'man' to find out more about commands, including arguments and parameters

# UNIX Operating System

- UNIX was initially developed by AT&T employees at Bell Labs in 1969
- Common Variants today: Linux, OpenSolaris, FreeBSD, HP/UX, AIX
- 100's of Major Linux Distributions: Debian-based (e.g. Ubuntu), Red Hat Package Manager (RPM)-based (SUSE, Red Hat & Fedora)
  - **Rajin uses CentOS**

# What is a Shell?

- The shell is a command line interpreter or shell that provides an interface to the UNIX operating system.
- A shell has a single purpose – to allow users to enter commands to execute, or create scripts containing commands, to direct the operation of the computer

```

Last login: Thu Apr 13 13:54:47 2017 from 202.7.176.94
#####
#           Welcome to the NCI National Facility!           #
#   This service is for authorised clients only. It is a criminal   #
#   offence to:                                           #
#       - Obtain access to data without permission           #
#       - Damage, delete, alter or insert data without permission #
#   Use of this system requires acceptance of the Conditions of Use #
#   published at http://nci.org.au/conditions/ #
#####
|   raijin.nci.org.au - 82408 processor InfiniBand x86_64 cluster |
|   User Guide:      http://help.nci.org.au   Assistance: help@nci.org.au |
|   Information:     http://nci.org.au |
|   Downtime Notices: http://nci.org.au/raijin-status/ |
=====

```

Jan 12 Additional Broadwell CPU Nodes Available  
 NCI has installed an additional 814 Broadwell-based compute nodes in Raijin. For instructions on the use of these nodes, please visit:  
<http://nci.org.au/broadwell>

Feb 16 User Job History  
 Users can now review completed jobs here:  
[https://usersupport.nci.org.au/report/job\\_history](https://usersupport.nci.org.au/report/job_history)

Mar 2 NVIDIA P100 Compute Nodes Online  
 NCI has installed new compute nodes with 4x NVIDIA P100 GPUs in each.  
 These are available under the 'gpupascal' queue.

To use these new nodes, your application would need to have been built against CUDA 8.0 (or newer). Details: <https://nci.org.au/gpu>

```

=====
[aw7523@raijin6 ~]$ 4$$

```

# More on UNIX Shells

- This course uses BASH
  - BASH = the Bourne-Again Shell
  - Default shell for most Linux systems
- Other common Shells include:
  - C Shell (csh)
  - K Shell (ksh)
  - TENEX C Shell (tcsh)

# What does BASH do?

- The shell (command line interpreter) interprets the commands entered by the user and passes those to the UNIX operating system.
- When enter a command and hit return
  - BASH parses the command line into tokens
  - 1<sup>st</sup> token is interpreted as the **command**
  - Remaining tokens are interpreted as **arguments**



# Anatomy of a command

- Commands comprise of:
  - a **command** that invokes a program
  - **arguments** to that program

`[user4@raijin ~]$ cmd arg1 arg2 arg3`

command prompt      command      Arguments (3)

`[user4@raijin ~]$ cmd arg1 arg2 arg3 "arg 4"`

username      Machine (login node)      command      Arguments (4)

# Commands and Shells

- Commands allow users to interact with the operating system via the *shell*
- Two-way communication is possible between the *shell* and commands
  - e.g. the `ls` command will return a list of files in a directory



# Exercise 1(a)

## Running Commands

Command	Description
<code>ls</code>	<i>Shows a directory listing</i>
<code>w</code>	Shows who is logged on to the system and what they are doing
<code>w iaa444</code>	Shows login, idle time, and what user iaa444 is doing
<code>finger</code>	Shows login, username & login details of users on the system
<code>finger iaa444</code>	Shows user iaa444's login details including name, idle time, login time as well as some of their environment settings
<code>date</code>	Prints the system time and date
<code>uptime</code>	Tells you how long the system has been running

# Command Line Options (Flags)

- **Command Line Options / Flags** modify the operation of the command.
  - There are two forms of flags – **Short Form & Long Form.**
  - Flags are case sensitive!
- **Short form options** start with a single hyphen “-”
  - `rm -f <filename>` → force removal of file
- **Long form Options** start with a double hyphen “--”
  - `rm --force <filename>` → force removal of file

**Both commands do the exact same thing!**

# Sample Command & Flags

- `ls -a`
  - short form flag “-a”
  - Shows all files incl. files starting with .
- `ls --all`
  - long form flag “--all”
  - Shows all files incl. files starting with .

# More Commands & Flags

– `ls -l`

- `short form flag "-l"`
- Shows the long format listing for all files in the directory

# Combining Flags

- Flags can be **combined in any order** but still mean the same thing
- The following command and flags all show the directory contents and show one file per line:

- `ls -a -l`
- `ls -l -a`
- `ls -la`
- `ls -al`
- `ls --all -l`
- `ls -l -all`



# Command Line Arguments

- **Command Line Parameters** are arguments sent to the program being called.
  - There are two forms of parameters – **Short Form & Long Form.**
  - Parameters are case sensitive!
- **Short form options** start with a single hyphen “-”
  - `tail -n 20 <filename>` → tail from the last 20 lines of the file
- **Long form Options** start with a double hyphen “--”
  - `tail --lines=20 <filename>` → tail from the last 20 lines of the file

# Exercise 1(b)

## Flags and Parameters

Command	Description
<code>ls</code>	Shows a directory listing
<code>ls -l</code>	(long) Lists directory contents of current directory using long listing format
<code>ls -a</code> <code>ls --all</code>	(all) Lists directory contents of current directory and does not ignore entries starting with .
<code>ls -la</code> <code>ls -all -a</code>	(all+long) Lists directory contents of current directory using long listing format and does not ignore entries starting with .
<code>ls -format=horizontal</code>	Lists directory contents of current directory horizontally
<code>ls --format=single-column</code>	Lists directory contents of current directory in a single column

# Using the “man” pages

- All Unix systems come equipped with manual or “man” pages which contain documentation about how each in-built command works
- Man pages can be accessed using the following command format:

**man** program, e.g. **man** ls

# The online manual

- **man** is a command that takes as its argument the name of another command.

## NAME

`find - search for files in a directory hierarchy`

## SYNOPSIS

`find [-H] [-L] [-P] [path...] [expression]`

- The [] indicate an optional argument (you don't type these)

# Exercise 1(c)

## The Calendar

Command	Description
<b>man</b> <i>&lt;command name&gt;</i>	Print the online manual entry for <i>&lt;command name&gt;</i>
<b>cal</b>	Print a calendar

# Unit 2: Where am I?

## Goals:

- Can explore navigate a hierarchy of directories
- Can specify files with relative and absolute paths
- Can create a hierarchy of directories
- Can copy and move files between directories

# UNIX Directories

- Called 'folders' under windows
  - Exactly the same concept
- You run into them a LOT more under Unix than you do under windows
- A directory is a container
  - It can contain files and other directories



What are the names of the contents of the directory called “/” (**root folder**)?

`/ac3`

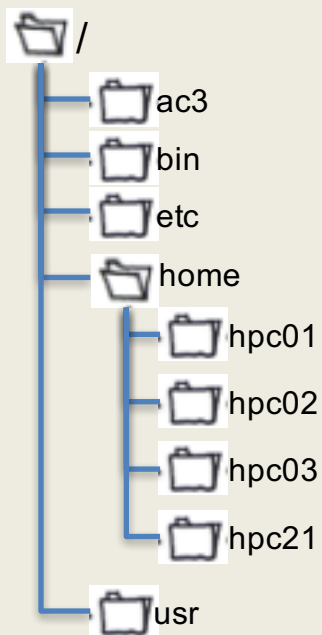
`/bin`

`/etc`

`/home`

`/usr`





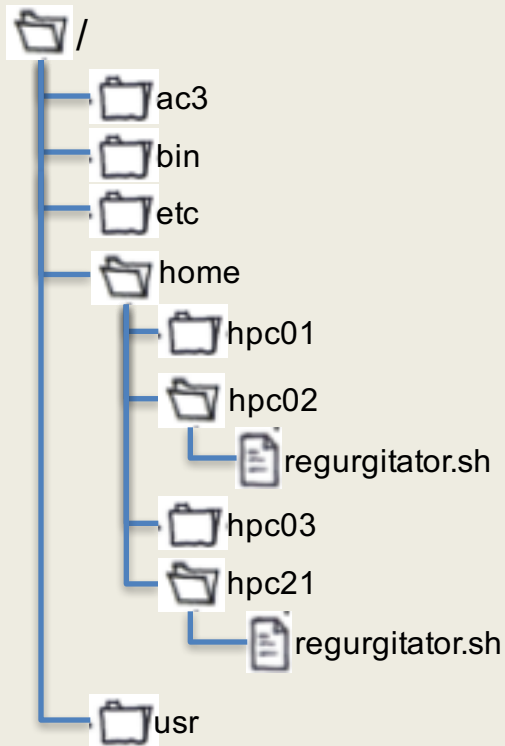
What are the names of the contents of the directory called **/home**?

**/home/hpc01**

**/home/hpc02**

**/home/hpc03**

**/home/hpc21**

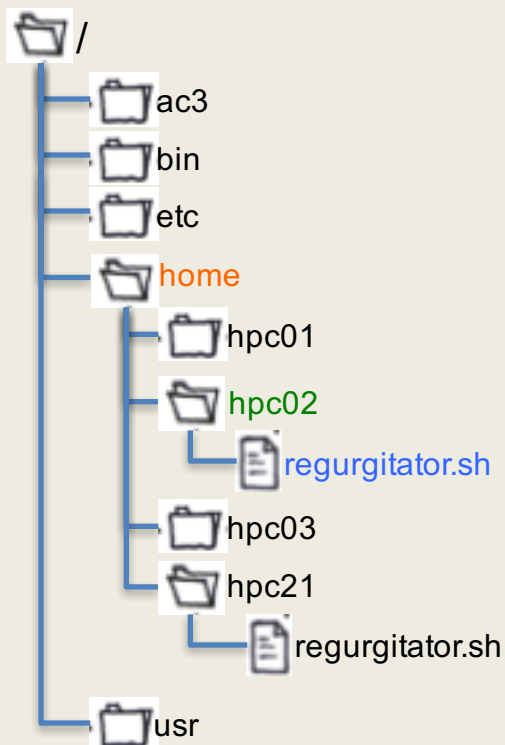


What are the names of the contents of the directory called **/home/hpc02**?

**`/home/hpc02/regurgitator.sh`**

What are the names of the contents of the directory called **/home/hpc21**?

**`/home/hpc21/regurgitator.sh`**



Path separator

`/home/hpc02/regurgitator.sh`

Path components

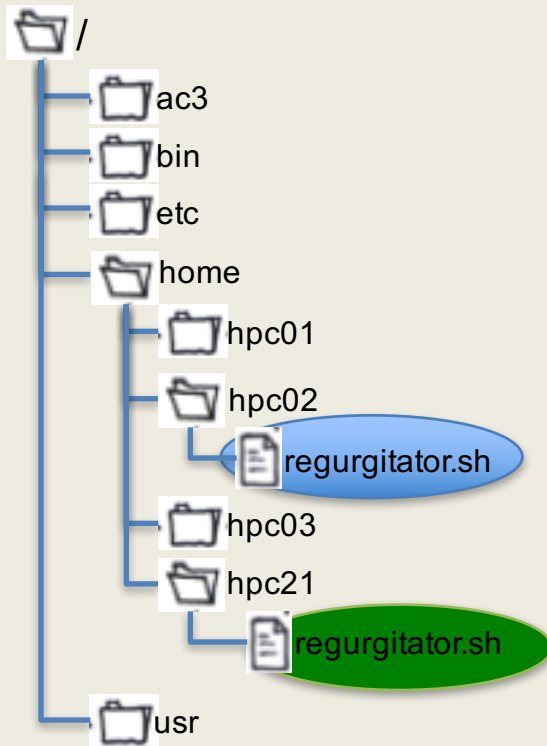
Where do I start from?

# Rules So Far

- There is one root, called “/”
  - This is different from windows, where there is one root for each disk drive C:, D:, etc
- A path from the root designates exactly one file: such a path is called an “absolute path”

# The Home Directory

- Each account (user) has a special directory called his/her home directory
  - That's where you 'get put' when you log in. (More on this later)
- BASH uses the "~" character to indicate "home directory"
- Two forms
  - ~ "my home directory"
  - ~**iaa444** "iaa444's home directory"



What are the names of the contents of the directory called ~

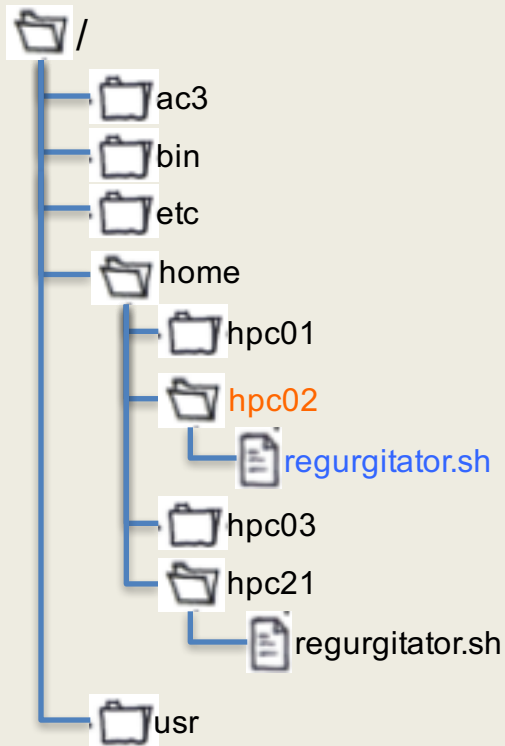
It depends who asks: If the user “hpc02” asks, then the answer is  
`/home/hpc02/regurgitator.sh`

What are the names of the contents of the directory called ~hpc21

`/home/hpc21/regurgitator.sh`

What is the absolute path of the file ~hpc02/regurgitator.sh

`/home/hpc02/regurgitator.sh`



Path separator

`~hpc02/regurgitator.sh`

Where do I start?

Path components

# Rules So Far

- There is one root, called “/”
- A path starting with “/” means “from the root”
- A path starting with “~” means “from the home”



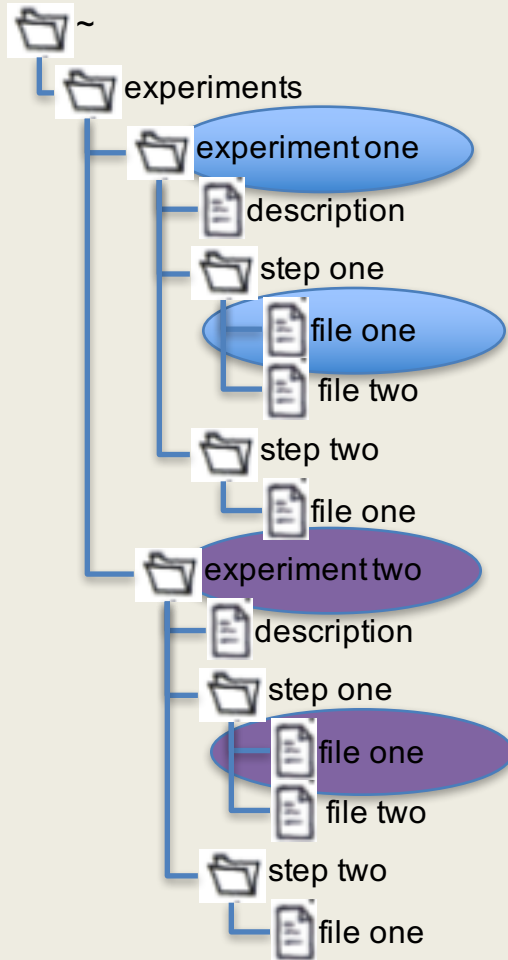
# Exercise 2(a)

## Finding your way around

Command	Description
<b>cd</b>	Change the working directory to your home directory
<b>cd</b> <i>&lt;path&gt;</i>	Change directory to <i>&lt;path&gt;</i>
<b>ls</b>	List the contents of the working directory
<b>ls</b> <i>&lt;path&gt;</i>	List the contents of <i>&lt;path&gt;</i>

# The Working Directory

- BASH maintains a “working” or “current” directory. That is “the directory that you are currently in”
  - Your prompt will show you your current working directory
- Linux uses the “.” character to denote the working directory. You can use this when running commands
  - `ls -la .` → show the listing for the current dir
- The command “**pwd**” will list the current working directory

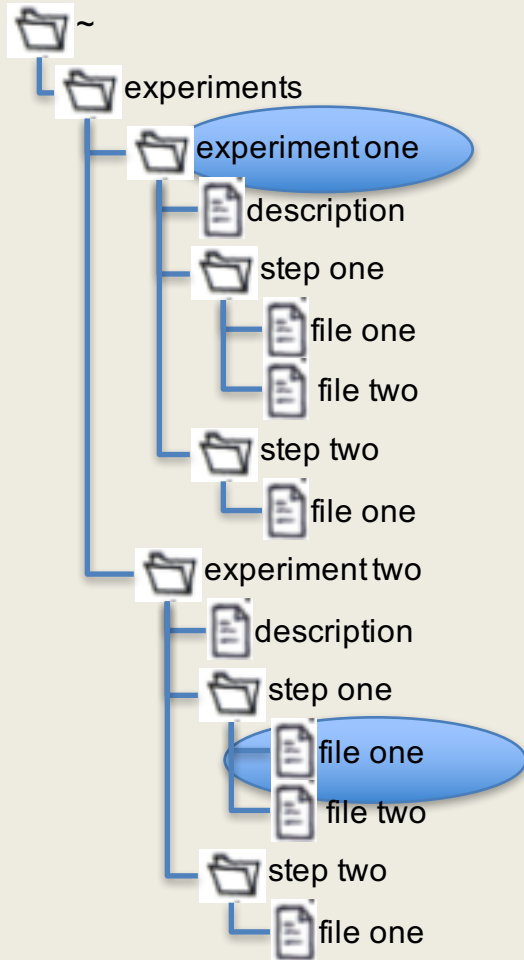


What is the relative path of “**step one/file one**”

What is the working directory?

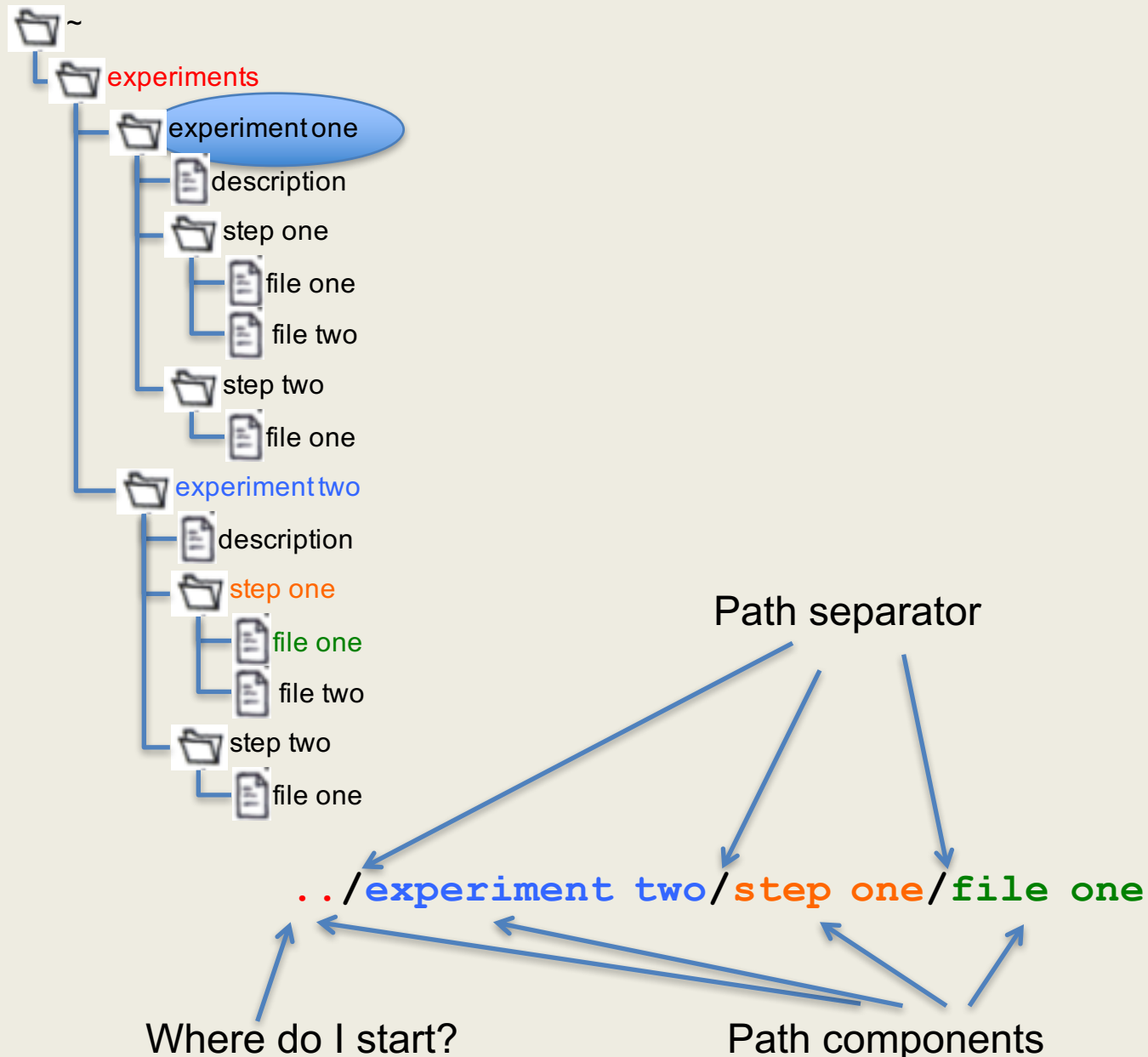
# Looking upwards

- UNIX uses “..” to denote the parent of a directory. You can use this when running commands, e.g.
  - `cd ..` → change up 1 directory
  - `cd ../hpc01` → change up 1 directory, then down 1 directory to hpc01
- UNIX understands “..” as a “normal directory” and it can appear in a path



What is the relative path of  
`../experiment two/step one/file one`

What is the working directory?



# Everything about directories

- There is one root, called “/”
- A path starting with “/” means “from the root”
- A path starting with “~” means “from the home”
- A path starting with anything else means “from the working directory”
- Note – the last and second last rules assume a hidden context (current user and current directory)

# Exercise 2(b)

## Finding your way around

Command	Description
<code>pwd</code>	Print the name of the current working directory
<code>cd ..</code>	Change directory to the parent directory
<code>ls</code>	List the contents of the working directory
<code>ls .</code>	List the contents of the working directory
<code>ls ..</code>	List the contents of the parent directory

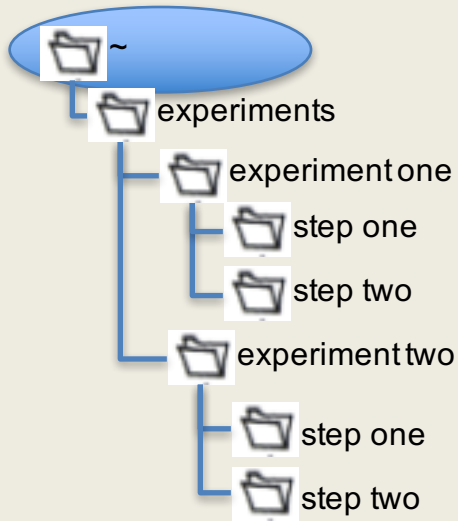


# Making New Directories

- You can create a directory, but only one at a time (by default)

```
mkdir </path/new_directory>
```

```
mkdir </path/to/new/directory>
```



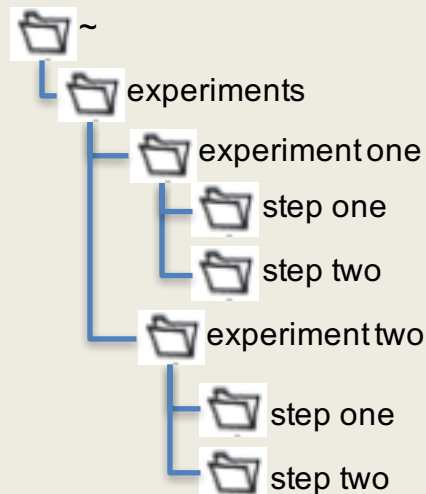
What happens when you execute the commands?

```
mkdir ~/experiments
cd ~/experiments
mkdir "experiment one"
mkdir "experiment one/step one"
mkdir "experiment one/step two"
mkdir "experiment two"
cd "experiment two"
mkdir "step one"
mkdir "step two"
```

# Exercise 2(c)

## Making Directories

Command	Description
<b>mkdir</b> <i>&lt;path&gt;</i>	Make a directory at <i>&lt;path&gt;</i>
<b>pwd</b>	Print the name of the current working directory
<b>cd</b> <i>&lt;path&gt;</i>	Change directory to <i>&lt;path&gt;</i>
<b>ls</b> <i>&lt;path&gt;</i>	List the contents of <i>&lt;path&gt;</i>

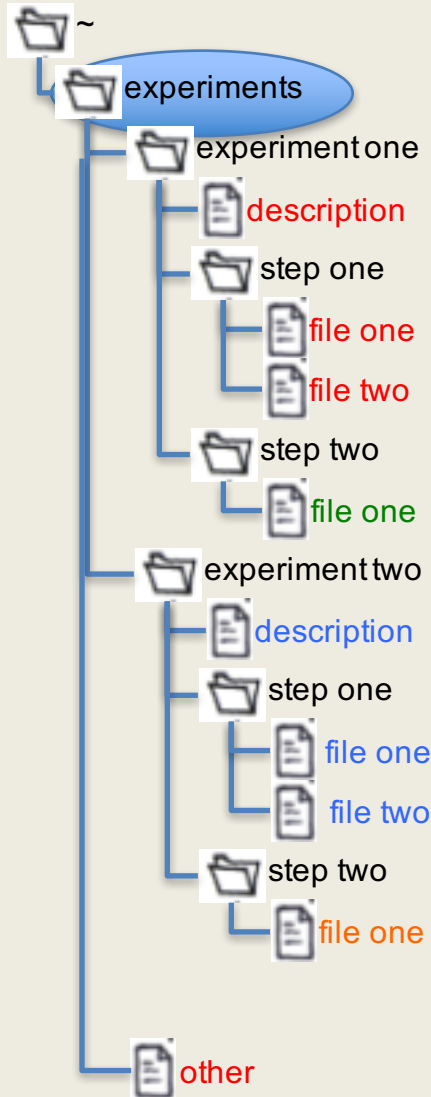


# Moving/Copying Files Around

- Files live in exactly one location
- Files can be copied and moved between directories

**cp** *<from path/file> <to path/file>*

**mv** *<from path/file> <to path/file>*



What happens when you execute the commands?

```
cp "experiment one/description" "experiment two"
```

```
cd "experiment one"
```

```
cp "step one/file one" "../experiment two/step two"
```

```
cp "../experiment two/description" ../other
```

```
cd "../experiment two/step one"
```

```
mv "file one" "../step two"
```

# Exercise 2(d)

## Moving and Copying Files

Command	Description
<b>cp</b> <i>&lt;from&gt;</i> <i>&lt;to&gt;</i>	Copy a file from <i>&lt;from&gt;</i> to <i>&lt;to&gt;</i>
<b>mv</b> <i>&lt;from&gt;</i> <i>&lt;to&gt;</i>	Copy a file from <i>&lt;from&gt;</i> to <i>&lt;to&gt;</i> then remove <i>&lt;from&gt;</i>
<b>man cp</b> <b>man mv</b>	Print the online manual entry for <i>&lt;command name&gt;</i> . What happens when you provide more than two arguments to <b>cp</b> and <b>mv</b> ?

# Unit 3: It's what's inside that counts

## Goals:

- Can look inside files (even really big ones)
- Can obtain rudimentary information about files

# What's in a file?

my dame has a lame tame crane  
my dame has a crane that is lame  
pray, gentle jane, that my dame's lame tame crane  
feed and come home again

0000000:	011011010111100100100000011001000110000101101101	my dam
0000006:	011001010010000001101000011000010111001100100000	e has
000000c:	011000010010000001101100011000010110110101100101	a lame
0000012:	001000000111010001100001011011010110010100100000	tame
0000018:	011000110111001001100001011011100110010100001010	crane.
000001e:	011011010111100100100000011001000110000101101101	my dam
0000024:	011001010010000001101000011000010111001100100000	e has
000002a:	011000010010000001100011011100100110000101101110	a cran
0000030:	011001010010000001110100011010000110000101110100	e that
0000036:	001000000110100101110011001000000110110001100001	is la
000003c:	011011010110010100001010011100000111001001100001	me.pra



# What's in a file?

This power-point presentation.

```
0000000: 01010000010010110000000110000010000010100000000000 PK....
0000006: 000001100000000000000100000000000000000000000000000 .....
```

000000c: 0010000100000000011010111110011110010000111101110 !...!.  
0000012: 0010001000000001100000000000000000000000010100011110 ".....  
0000018: 0000000000000000000000010011000000000000100000000010 .....

000001e: 010110110100001101101111011011100111010001100101 [Conte  
0000024: 011011100111010001011111010101000111100101110000 nt\_Typ  
000002a: 011001010111001101011101001011100111100001101101 es].xm  
0000030: 011011000010000010100010000001000000001000101000 l ... (  
0000036: 101000000000000000000000100000000000000000000000000 .....

000003c: 000

# What is a 'text file'?

109

104

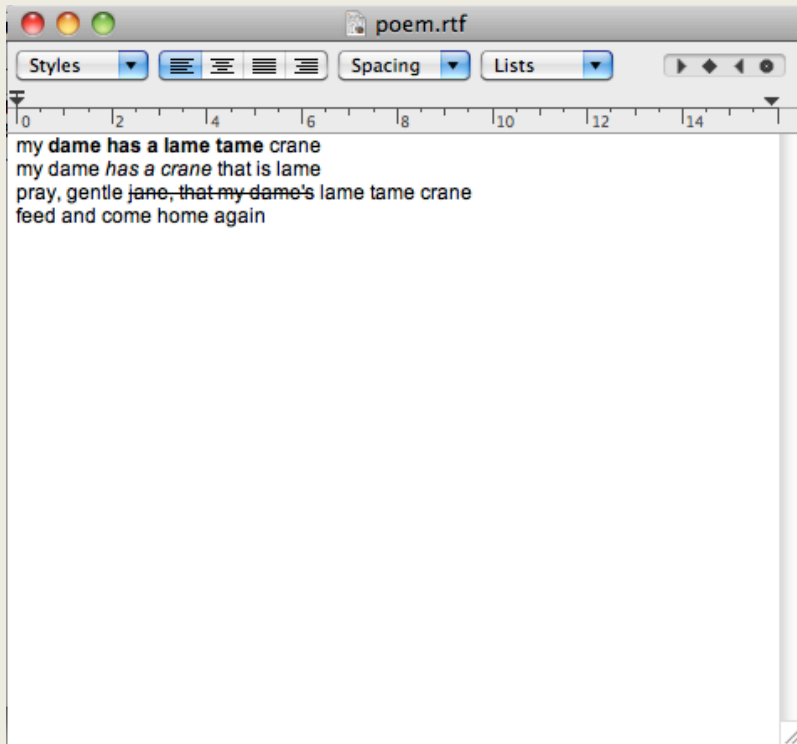
10

0000000:	<b>01101101</b> 0111100100100000011001000110000101101101	my dam
0000006:	0110010100100000 <b>01101000</b> 011000010111001100100000	e <b>h</b> as
000000c:	011000010010000001101100011000010110110101100101	a lame
0000012:	001000000111010001100001011011010110010100100000	tame
0000018:	0110001101110010011000010110111001100101 <b>00001010</b>	crane.
000001e:	011011010111100100100000011001000110000101101101	my dam
0000024:	011001010010000001101000011000010111001100100000	e has
000002a:	011000010010000001100011011100100110000101101110	a cran
0000030:	011001010010000001110100011010000110000101110100	e that
0000036:	001000000110100101110011001000000110110001100001	is la
000003c:	011011010110010100001010011100000111001001100001	me.pra

Dec	Hx	Oct	Char	Dec	Hx	Oct	Html	Chr	Dec	Hx	Oct	Html	Chr	Dec	Hx	Oct	Html	Chr
0	0	000	<b>NUL</b> (null)	32	20	040	&#32;	<b>Space</b>	64	40	100	&#64;	<b>@</b>	96	60	140	&#96;	<b>`</b>
1	1	001	<b>SOH</b> (start of heading)	33	21	041	&#33;	<b>!</b>	65	41	101	&#65;	<b>A</b>	97	61	141	&#97;	<b>a</b>
2	2	002	<b>STX</b> (start of text)	34	22	042	&#34;	<b>"</b>	66	42	102	&#66;	<b>B</b>	98	62	142	&#98;	<b>b</b>
3	3	003	<b>ETX</b> (end of text)	35	23	043	&#35;	<b>#</b>	67	43	103	&#67;	<b>C</b>	99	63	143	&#99;	<b>c</b>
4	4	004	<b>EOT</b> (end of transmission)	36	24	044	&#36;	<b>\$</b>	68	44	104	&#68;	<b>D</b>	100	64	144	&#100;	<b>d</b>
5	5	005	<b>ENQ</b> (enquiry)	37	25	045	&#37;	<b>%</b>	69	45	105	&#69;	<b>E</b>	101	65	145	&#101;	<b>e</b>
6	6	006	<b>ACK</b> (acknowledge)	38	26	046	&#38;	<b>&amp;</b>	70	46	106	&#70;	<b>F</b>	102	66	146	&#102;	<b>f</b>
7	7	007	<b>BEL</b> (bell)	39	27	047	&#39;	<b>'</b>	71	47	107	&#71;	<b>G</b>	103	67	147	&#103;	<b>g</b>
8	8	010	<b>BS</b> (backspace)	40	28	050	&#40;	<b>(</b>	72	48	110	&#72;	<b>H</b>	104	68	150	&#104;	<b>h</b>
9	9	011	<b>TAB</b> (horizontal tab)	41	29	051	&#41;	<b>)</b>	73	49	111	&#73;	<b>I</b>	105	69	151	&#105;	<b>i</b>
10	A	012	<b>LF</b> (NL line feed, new line)	42	2A	052	&#42;	<b>*</b>	74	4A	112	&#74;	<b>J</b>	106	6A	152	&#106;	<b>j</b>
11	B	013	<b>VT</b> (vertical tab)	43	2B	053	&#43;	<b>+</b>	75	4B	113	&#75;	<b>K</b>	107	6B	153	&#107;	<b>k</b>
12	C	014	<b>FF</b> (NP form feed, new page)	44	2C	054	&#44;	<b>,</b>	76	4C	114	&#76;	<b>L</b>	108	6C	154	&#108;	<b>l</b>
13	D	015	<b>CR</b> (carriage return)	45	2D	055	&#45;	<b>-</b>	77	4D	115	&#77;	<b>M</b>	109	6D	155	&#109;	<b>m</b>
14	E	016	<b>SO</b> (shift out)	46	2E	056	&#46;	<b>.</b>	78	4E	116	&#78;	<b>N</b>	110	6E	156	&#110;	<b>n</b>
15	F	017	<b>SI</b> (shift in)	47	2F	057	&#47;	<b>/</b>	79	4F	117	&#79;	<b>O</b>	111	6F	157	&#111;	<b>o</b>
16	10	020	<b>DLE</b> (data link escape)	48	30	060	&#48;	<b>0</b>	80	50	120	&#80;	<b>P</b>	112	70	160	&#112;	<b>p</b>
17	11	021	<b>DC1</b> (device control 1)	49	31	061	&#49;	<b>1</b>	81	51	121	&#81;	<b>Q</b>	113	71	161	&#113;	<b>q</b>
18	12	022	<b>DC2</b> (device control 2)	50	32	062	&#50;	<b>2</b>	82	52	122	&#82;	<b>R</b>	114	72	162	&#114;	<b>r</b>
19	13	023	<b>DC3</b> (device control 3)	51	33	063	&#51;	<b>3</b>	83	53	123	&#83;	<b>S</b>	115	73	163	&#115;	<b>s</b>
20	14	024	<b>DC4</b> (device control 4)	52	34	064	&#52;	<b>4</b>	84	54	124	&#84;	<b>T</b>	116	74	164	&#116;	<b>t</b>
21	15	025	<b>NAK</b> (negative acknowledge)	53	35	065	&#53;	<b>5</b>	85	55	125	&#85;	<b>U</b>	117	75	165	&#117;	<b>u</b>
22	16	026	<b>SYN</b> (synchronous idle)	54	36	066	&#54;	<b>6</b>	86	56	126	&#86;	<b>V</b>	118	76	166	&#118;	<b>v</b>
23	17	027	<b>ETB</b> (end of trans. block)	55	37	067	&#55;	<b>7</b>	87	57	127	&#87;	<b>W</b>	119	77	167	&#119;	<b>w</b>
24	18	030	<b>CAN</b> (cancel)	56	38	070	&#56;	<b>8</b>	88	58	130	&#88;	<b>X</b>	120	78	170	&#120;	<b>x</b>
25	19	031	<b>EM</b> (end of medium)	57	39	071	&#57;	<b>9</b>	89	59	131	&#89;	<b>Y</b>	121	79	171	&#121;	<b>y</b>
26	1A	032	<b>SUB</b> (substitute)	58	3A	072	&#58;	<b>:</b>	90	5A	132	&#90;	<b>Z</b>	122	7A	172	&#122;	<b>z</b>
27	1B	033	<b>ESC</b> (escape)	59	3B	073	&#59;	<b>;</b>	91	5B	133	&#91;	<b>[</b>	123	7B	173	&#123;	<b>{</b>
28	1C	034	<b>FS</b> (file separator)	60	3C	074	&#60;	<b>&lt;</b>	92	5C	134	&#92;	<b>\</b>	124	7C	174	&#124;	<b> </b>
29	1D	035	<b>GS</b> (group separator)	61	3D	075	&#61;	<b>=</b>	93	5D	135	&#93;	<b>]</b>	125	7D	175	&#125;	<b>}</b>
30	1E	036	<b>RS</b> (record separator)	62	3E	076	&#62;	<b>&gt;</b>	94	5E	136	&#94;	<b>^</b>	126	7E	176	&#126;	<b>~</b>
31	1F	037	<b>US</b> (unit separator)	63	3F	077	&#63;	<b>?</b>	95	5F	137	&#95;	<b>_</b>	127	7F	177	&#127;	<b>DEL</b>

Source: [www.LookupTables.com](http://www.LookupTables.com)

# Two sides of the same coin



```
{\rtf1\ansi\ansicpg1252\cocoartf1038\cocoa  
subrtf350
```

```
{\fonttbl{\f0\fswiss\fcharset0 ArialMT;}
```

```
{\colortbl;\red255\green255\blue255;}
```

```
\paperw11900\paperh16840\margl1440\ma  
rgr1440\vieww9000\viewh8400\viewkind0  
\deftab720
```

```
\pard\pardeftab720\ql\qnatural
```

```
\f0\fs24 \cf0 my
```

```
\b dame has a lame tame
```

```
\b0 crane\
```

```
my dame
```

```
\i has a crane
```

```
\i0 that is lame\
```

```
pray, gentle \strike \strikec0 jane, that my  
dame's\strike0\striked0 lame tame crane\
```

```
feed and come home again\
```

```
}
```

# Why do we care?

- BASH is more or less committed to ASCII text
- In particular
  - your keystrokes result in ASCII codes being sent to the shell
  - when a program generates output, it is interpreted by the shell as ASCII codes and displayed on the screen (even some non-printables)

# Exercise 3(a)

## Looking into files

Command	Description
<b>ls</b>	You've met before
<b>cat</b> <filename>	Print a file to the terminal ( <b>cat</b> enate)
<b>less</b> <filename>	Like <b>cat</b> , but <b>less</b> at a time
<b>head</b> <filename>	Like <b>cat</b> , but just the top of the file
<b>tail</b> <filename>	Like <b>cat</b> , but just the end of the file
<b>wc</b> <filename>	<b>W</b> ord <b>c</b> ount – count the words in a file
<b>du</b>	<b>D</b> isk <b>u</b> sage – how much space is used

# Unit 4: It's all just files

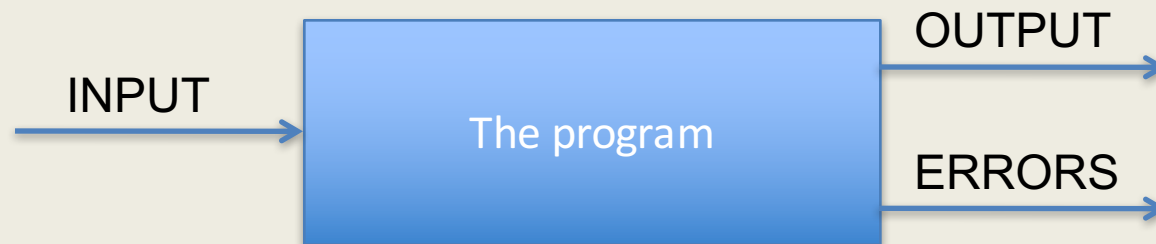
## Goals:

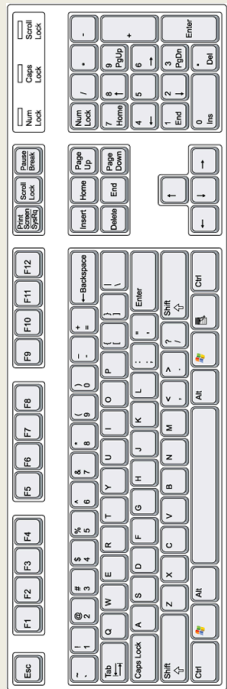
- Can save the output of programs into files
- Can chain commands together

# The core magic of the Linux command line

- The next bit is the core magic of Linux, as far as this module goes.
  - The keyboard is just a file
  - The console is just a file
- Define – What is input?
- Define – What is output?







INPUT

The program

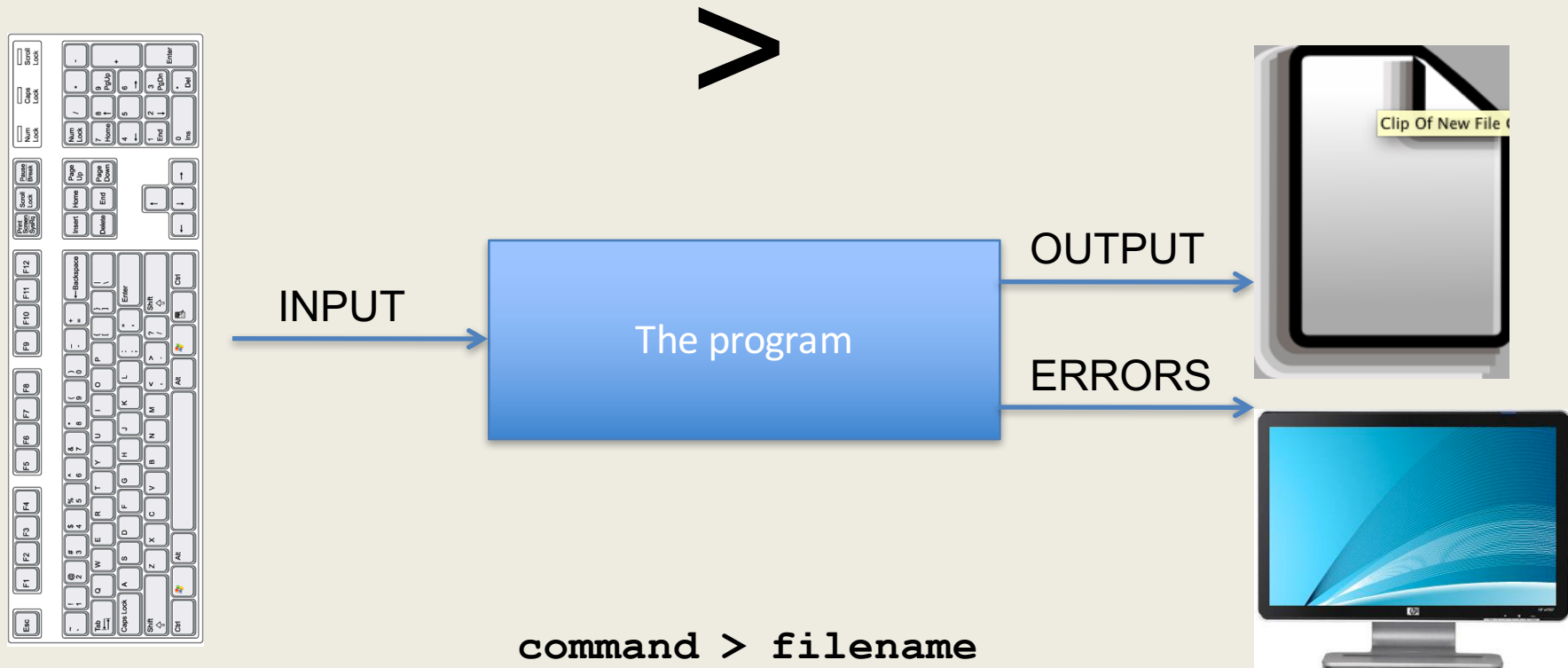
OUTPUT

ERRORS



command

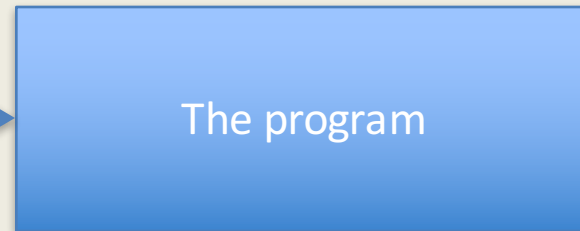
# Redirect OUTPUT elsewhere...



# Append OUTPUT elsewhere...

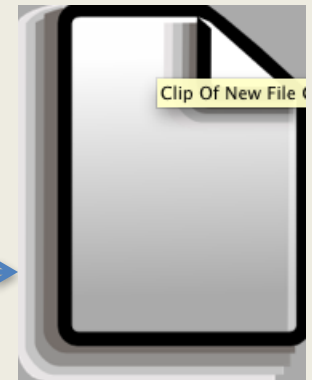


INPUT



OUTPUT

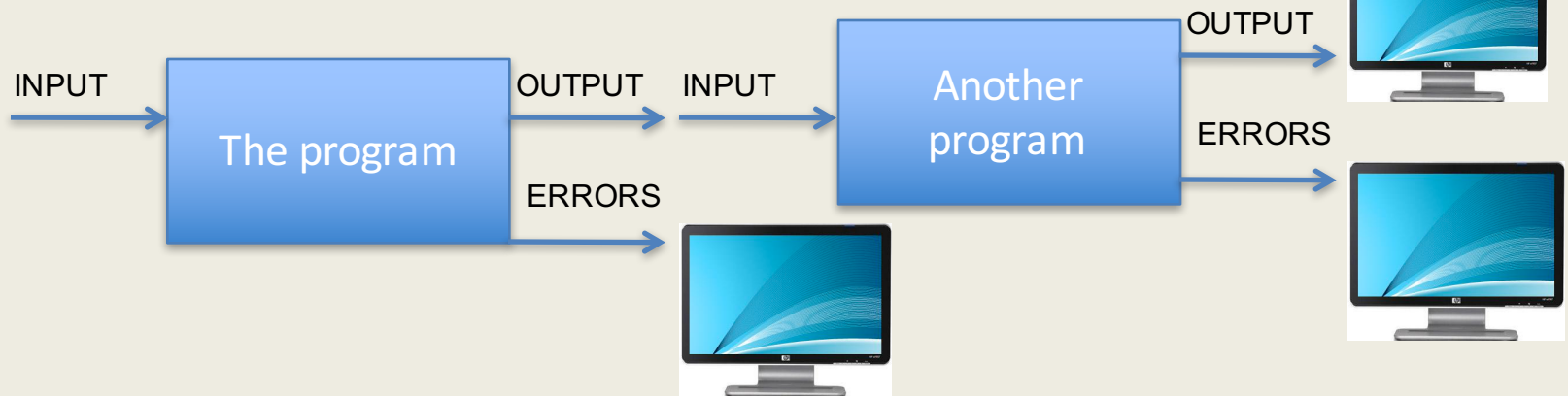
ERRORS



`command >> filename`

# Pipe OUTPUT to a command...

|



`command | another_command`

# Exercise 4(a)

## Looking into files

Command	Description
<b>fortune</b>	Tells your fortune – really! Try it!
<b>sort</b> <i>&lt;file&gt;</i>	Sorts the contents of <i>&lt;file&gt;</i> , or input if no file is specified.
<b>uniq</b> <i>&lt;file&gt;</i>	Removes any redundant lines in <i>&lt;file&gt;</i> , or input if no file is specified. Redundant lines have to be adjacent to be considered redundant.
<b>cat</b> <i>&lt;file&gt;</i>	Prints the contents of <i>&lt;file&gt;</i>
<i>&lt;command&gt;</i> <b>&gt;</b> <i>&lt;file&gt;</i>	Run <i>&lt;command&gt;</i> but rather than print the output to the console, <b>redirect the output to</b> <i>&lt;file&gt;</i> .
<i>&lt;command&gt;</i> <b>&gt;&gt;</b> <i>&lt;file&gt;</i>	Run <i>&lt;command&gt;</i> but rather than print the output to the console, <b>append the output to</b> <i>&lt;file&gt;</i> .
<i>&lt;command&gt;</i> <b> </b> <i>&lt;command&gt;</i>	Pipe one <i>&lt;command&gt;</i> to another <i>&lt;command&gt;</i>

# Unit 5: Products of our environment

## Goals:

- Can set and read environment variables
- Can add programs to the PATH
- Can query the environment for which variables exist

# Case study – finding programs

- Recall from previous units
  - Sometimes you have to specify the path of a program

```
regurgitator.sh
```

```
./regurgitator.sh
```

```
../regurgitator.sh
```

```
/bin/regurgitator.sh
```



# PATH

- There is a 'special' thing called the PATH
  - Part of what makes it special is encoded
  - Part of what makes it special is convention
- There is an environment full of variables
  - One of those variables is called PATH
  - Use '\$' to tell BASH you're talking about a variable
- To check what an environment variable is set to, use `$<ENVIRONMENT_VARIABLE>`
  - e.g. `$PATH`

# A sample path

```
/bin:/usr/bin:/usr/local/bin:.
```

# What does BASH do?

- When you hit return
  - BASH parses the command line into tokens
  - Any 'special' tokens are expanded
  - The first token is treated as the command
  - The remaining tokens are passed to the command as arguments

# A simple example

- Your home directory is stored in the environment, you can be taken there directly

```
cd $HOME
```

# Exercise 5(a)

## Realistic Evaluation of the Environment

Command	Description
<code>\$&lt;VARIABLE&gt;</code>	Evaluates to the contents of the variable <code>&lt;VARIABLE&gt;</code> .
<code>env</code>	Prints all environment variables and their values.
<code>echo &lt;arg1&gt; &lt;arg2&gt; ...</code>	Prints out its arguments, with no additional information.
<code>grep &lt;pattern&gt; &lt;file&gt;</code>	<p>Prints out all occurrences of <code>&lt;pattern&gt;</code> in <code>&lt;file&gt;</code>. If no file is specified, processes standard input. For example</p> <p style="text-align: center;"><code>env   grep PA</code></p> <p>will print all variables containing 'PA' in them or their value.</p>

# Manipulating the Environment

- You can use the environment to your advantage.
  - Update the path
  - Change your prompt
  - Use the environment to share information between your scripts

# The export command

The **export** command is used to set the value of an environment variable:

```
export  <VARIABLE_NAME>=<VALUE>
```

# Exercise 5(b)

## Manipulating the Environment

Command	Description
<b>export</b> <i>&lt;VAR&gt;=&lt;VAL&gt;</i>	Set a variable called <i>VAR</i> and give it contents <i>VAL</i> . The variable will be available to programs invoked from the shell.



# Summing up – Your Skills

- You have now got the rudimentary skills to move around, make and view files, chain commands together and manipulate the environment

# Summing up – Hard Lessons

- BASH is wildly picky
- Silent success
- Command-line work is flexible but error prone
- There's a lot of hidden context
  - ENVIRONMENT, WORKING DIRECTORY, STDOUT/STDIN, COMMAND LINE EXPANSION
- BASH is wildly picky

# Summing up – the positives

- Command line gives you a more **powerful** means to control a program
- Allows **direct interaction** with shell commands (no GUI needed)
- Allow for **ease of automation** via scripting
- UNIX is more **flexible** and can be installed on many different types of machines (e.g. mainframes/supercomputers)
- Unix possesses much greater **processing power** than Windows.