CITS4407
Open Source Tools and Scripting

Michael J. Wise

Lecture 0
Introduction

Who we are

- Assoc Prof Michael Wise – Coordinator, lecturer

- Aline Andrade – Lab Facilitator

- Bryce O'Connor – Lab Facilitator

- Jiachuan Liu – Lab Facilitator

What is CITS4407 About?

- Open Source philosophy/Unix philosophy
  - *Open source tools, and tool ensembles (pipelines), power much of the world, particularly internet and scientific computing*
- Creating and using shell scripting to combine software tools for data cleaning, first pass data analysis, rapid prototyping

Advisable Prior Study

- There are no formal perquisites for CITS4407

- However, if you have not done a prior programming unit before, such as CITS1401, the unit may be a stretch

  - *Computational Thinking*

Why Bother?

- Automate repetitive tasks
  - *Print 1 file using the GUI, no problem*
  - *Print 200 files ….*
- Rapid prototyping
  - *A quick and dirty solution right now may be all you need*

Demo

- I have a text file, called `Alice_in_Wonderland.txt` that contains the text of the book "Alice in Wonderland", Lewis Carrol. (Guttenberg Project)

    - *You can find it next to the Lecture Slides PDF for Lecture 0 on LMS*

- I want to extract all the words (just sequences of letters), count the number of times each unique word appears, and then list the words in descending order of occurrence.

- Get together with others and come up with an estimate how long it would take to code up a solution.

Course Outcomes

- Understand the Open Source, and in particular Unix, philosophy

- Understand what shell scripting is suited for, and what it's not that well suited for

- Confidently use a number of the common Unix/Linux tools

- Be able to write Bash/Shell Tools scripts to:

  - *Solve small problems*

  - *Automate repetitive computational tasks*

Bash as a Programming Language

- We don't assume you know a programming language coming into this unit, but knowing something of Python, Java, C, etc, will help

- Bash is the most used of the Unix shells, found on all Linux machines, Mac OSX (zsh is largely Bash, but not the same).

- Why Shell rather than, say, Python/Java/C?

  – *Shell scripting very good for quickly writing "glueware"*

Textbook, Web page and Resources

- There is no set text for this unit. One useful free (creative commons license) ebook is, "The Linux Command Line: A Complete Introduction" (5e), William E. Shotts Jr, 2019 http://linuxcommand.org/tlcl.php

- The Awk material is covered by "The GNU Awk User's Guide (5e)", https://www.gnu.org/software/gawk/manual/gawk.html

- Everything you need can be found on LMS

- There is a Resources tab on the unit web page

Implementations

- Laptop only, not tablets (or phones)

- Linux

  - *Ubuntu preferred, V 22.04. Bash is generally the default shell*

- Mac OSX

  - *Terminal.app gives you zsh, which is close to Bash, but not identical. Some Unix commands also slightly different*

    - FreeBSD versus GNU

  - *Better to install and use Docker, or UniApps.*

- Windows

  - *Install and use Docker on top of Windows Subsystem for Linux (WSL) , or UniApps*

Implementations

- UWA recently released UniApps

- UniApps includes Linux Lab, which can also be used for this unit

The standard environment expected for this unit is:
Ubuntu Linux with the Bash shell and GNU tool set.

Code developed on other systems may not work on the standard system, so at least check before submission.

Organisation

- 2 x 1hr lectures a week
  - *A Lecture may take more than 1 slot*
  - *Slots are, in fact 45 mins, starting on the hour*
  - *If you can, bring your laptop*

- 1 programming lab per week (2 hrs)
  - *Lab demonstrators available*
  - *Starts Week 2 (watch out for public holidays)*
  - *Check your Timetable/CAS; multiple time slots across the week*

Labs - Expectations

- Labs are not assessed, but if you want to do well in the unit you should attend at least one lab session per week starting in Week 2

  - *Some learning in the unit will only take place in labs*

- You are welcome to attend as many lab sessions as you want

  - *preference to those timetabled to be there*

- Please bring your own laptop with Bash/Unix Tools/Docker and/or UniApps installed.

- **This is your time to work on relevant exercises from worksheets with help at hand**

Assessment

- Assessment is based on both
  - *Understanding of fundamental concepts*
  - *Practical problem-solving and programming skills*
- Two programming projects
  - *Assignment 1 due Mon. of Week 8 (worth 20%)*
  - *Assignment 2 due Mon. of Week 12 (worth 20%)*
- Final Exam
  - *2hr face-to-face (single page of notes) exam in the exams period (60%)*
- Summative online assessment (marked, but not for marks). Week 7
- See Weekly Schedule

Getting Help

- Labs

- Materials in Resources

- Discussion Forum on LMS

- Above all, seek help early.

Svengraph, WikiMedia

Do Something Useful in Week 1

- If you are new to UWA
  - *Get your computer account name and password*
  - *Organize your UWA email account*
  - *Find out which lab you're in*
- Check that you have Ubuntu 20.04+ (Linux), or install Docker(OSX), WSL + Docker (Windows)
  - *Ubuntu 20.04 also fine. Make sure it includes gawk*
  - *Will happen in Lab 1 in Week 2*
  - *UniApps*

Other Stuff

- Interesting Things Page! (Prize every so often for best contribution – as judged by me 😉)

- Prize for any errors detected!

- I have set slides in Century Schoolbook font (with some Courier and Arial for computer code and meta–language). If you have trouble reading it, please let me know

  – *Accessibility is important*

Other Stuff

- "10 Signs You Will Suck at Programming"
  - *Article linked to Interesting Things page*
  - *Has really great advice about what you need to succeed at programming*
  - *READ IT*

- Engage with the unit!!!
  - *Good data to show that if you turn up to lectures and generally engage with the unit, you will do better (Drouin, 2014, Edwards &Clinton 2018) – see Interesting Things.*

- Have fun!!!

Acknowledgements

- It is important to acknowledge the earlier work on this unit that was done by Chris MacDonald, and more recently, Arran Stewart and Daniel Smith

PheobeA - Redbubble