Research-led teaching in INFO1903

Tara Murphy and James R. Curran

{tm,james}@it.usyd.edu.au

The University of Sydney

In 2006, we developed and ran a new first year course, INFO1903 Foundations of Information Technology (Adv). One of the aims of this course was to teach advanced science students the IT skills they will need in research. Research-led teaching was a critical component of the course.

Introduction

IT is pervasive in modern science. IT skills are critical for researchers in any domain, from physics, chemistry, biology to economics and architecture. Many of our top students will go on to do honours and PhDs or work in domains where IT skills are essential. In this context, the aims of INFO1903 are to:

- teach science students the IT skills they need to excel in their scientific research
- give IT students a broad overview of computer science and IT in the context of real data and real world problems.

In both cases we try to impart a philosophy of expert use of a wide range of tools and a critical approach to solving new problems with these tools. We also try to place IT in a broader social and historical context – a large part of which is its role in research. One of the best ways of doing this was to incorporate real research data throughout the course, in addition to teaching research-related skills.

Course Outline

INFO1903 is an ambitious course, targeted at our top students. It covers a wide range of topics and skills in IT including:

- Advanced use of productivity tools (Word, Excel)
- Unix tools and shell scripting
- Programming in Python
- Databases and SQL
- Internet technologies (HTML, CSS, XML, web services)
- Presentations and report writing
- Ethics in IT

There were three lectures and three hours of lab work a week. Programming was taught for one lecture a week throughout the semester.

Incorporating Research Data

As far as possible, all tutorial exercises, assignments, and the major project were based on real research data. In particular we drew examples from astronomy and computational linguistics – our areas of expertise. We also used biology, stock market and web log data.

The early assignments required students to solve data processing problems in Unix/Python on a variety of datasets. For example:

A search of the NASA Extragalactic Database for objects near the position RA=14h23m34.00000s, Dec=+0d00m21.0000s resulted in the file you have been given. We want you to extract some information from this HTML file and load it into Excel for further processing...

Major Project

Throughout the semester they then worked on arranging the data into databases, analysing it in Excel and writing CGI scripts to access it from a website. The final project consisted of setting up a backend and frontend to collection of linguistics data so that it could be easily queried by researchers. An example of a frontend to this system is shown below.
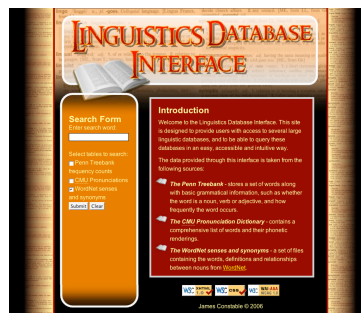
Communicating Research

We also taught specific skills that we think would be useful for students going on into scientific research projects:

- building websites for querying and presenting data
- data visualisation and presentation
- writing a scientific report (in Word and LaTeX)
- giving a presentation for a scientific conference

BOTE

An essential skill for scientists and IT professionals is the ability to estimate or perform **Back of the Envelope** calculations. Ten minutes in the middle of each lecture was spent like this:
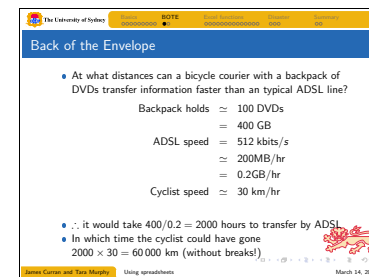
Student Feedback

The feedback we received from the first group of students to do the course was very positive. Students responded well to the range of domains covered in the examples. Some of the responses to the ITL survey which are relevant to research-led teaching approach are:

- *"I definitely gained some valuable skills from this subject that I will find useful later in my studies, especially regarding the use of Unix tools, MySQL and CGI scripting"*
- *"Two assessments involved scientific report writing and speech making which are good skills to have even outside IT"*
- *"This course allows for the development of these type of generic skills where most other courses do not. This is a very valuable aspect of this course and I highly commend it for that."*
- *"...the practical things we learnt are obviously applicable and useful in real situations, its been a really enjoyable course."*

Future Development

For 2007 we want to do the following:

- Expand to larger number of students
- Expand the range of domains that examples are drawn from
- Incorporating an automatic marking system for Python
- Try using SMS for lecture theatre voting

http://www.it.usyd.edu.au/~info1903