

## Aims

Get used to more Perl features.

## Assessment

**Submission:** give cs2041 lab06 whale.pl all\_whales.pl prereq.pl [recursive\_prereq.pl digits.py echon.py simple\_tail.py whale.py all\_whales.py]

**Deadline:** either during the lab, or Monday 5 September 11:59pm (midnight)

**Assessment:** Make sure that you are familiar with the lab assessment criteria (<lab/assessment.html>).

## Background

We have covered only a small amount of Perl in lectures. In fact, to cover the whole language in detail would take a whole semester, so we're going to rely on you finding out about the language yourself in tutes, labs and assignments. A good place to start is the Perl documentation & tutorial links on the class home page. For example you might find these useful:

- Perl language syntax (<http://search.cpan.org/dist/perl/pod/perlsyn.pod>)
- Perl functions (<http://search.cpan.org/dist/perl/pod/perlsub.pod>)
- Perl operators (<http://search.cpan.org/dist/perl/pod/perlop.pod>)

## Input Data

A biologist needs your help counting migrating whales. As pods of migrating whales as they swim past the biologist's boat, the biologist types the number and species of whales in the pod.

Here is an example of the input the biologist will provide

```
15 humpback
2 orca
2 sperm whale
19 beluga
2 humpback
2 sperm whale
4 orca
4 pygmy right whale
7 humpback
1 orca
```

## Counting One Whale Species

Write a Perl script `whale.pl` which given a whale name as a command line argument, reads whale observations in the above format until the end-of-input is reached and then prints the number of pods of the specified whale and the total numbers of whales in those pods. For example:

```
$ ./whale.pl 'sperm whale'
15 humpback
2 orca
2 sperm whale
19 beluga
2 humpback
2 sperm whale
4 orca
4 pygmy right whale
7 humpback
1 orca
Ctrl-d
sperm whale observations: 2 pods, 4 individuals
```

```
$ ./whale.pl "cuvier's beaked whale"
```

```
3 orca
```

```
673 beluga
```

```
2 sperm whale
```

```
123 pilot whale
```

```
2 cuvier's beaked whale
```

```
19 beluga
```

```
2 north atlantic right whale
```

```
2 humpback
```

```
5 cuvier's beaked whale
```

```
2 sperm whale
```

```
1 north atlantic right whale
```

```
1119 beluga
```

```
7 orca
```

```
4 pygmy right whale
```

```
7 humpback
```

```
1 false orca
```

```
3 beluga
```

```
4 pilot whale
```

```
1 orca
```

```
Ctrl-d
```

```
cuvier's beaked whale observations: 2 pods, 7 individuals
```

As usual use `autotest` to assist in testing your code and push your work to `gitlab.cse.unsw.edu.au` every time you make some progress with it.

```
$ ~cs2041/bin/autotest lab06 whale.pl
```

```
...
```

```
$ git add whale.pl
```

```
$ git commit -a -m "whales are cool just like Andrew"
```

```
...
```

```
$ git push
```

```
...
```

## Counting All Whale Species

Write a Perl script `all_whales.pl` that reads whale observations in the above format until the end-of-input is reached and for all species of whale observed prints the number of pods of the specified whale and the total numbers of whales in those pods. For example:

```

$ ./all_whales.pl
15 humpback
2 orcas
2 sperm whales
19 belugas
2 humpbacks
2 sperm whales
4 Orcas
4 pygmy right whale
7 humpbacks
1 orca
3 ORCAS
673 belugas
2 sperm whales
123 pilot whale
2 cuvier's beaked whales
19 beluga
2 north atlantic right whale
2 humpbacks
5 cuvier's BEAKED whale
2 Sperm Whales
1 north atlantic right whales
1119 belugas
7 orcas
4 pygmy right whales
7 humpbacks
1 False Orca
3 belugas
4 pilot whaleS
1 orca
Ctrl-d
beluga observations: 5 pods, 1833 individuals
cuvier's beaked whale observations: 2 pods, 7 individuals
false orca observations: 1 pods, 1 individuals
humpback observations: 5 pods, 33 individuals
north atlantic right whale observations: 2 pods, 3 individuals
orca observations: 6 pods, 18 individuals
pilot whale observations: 2 pods, 127 individuals
pygmy right whale observations: 2 pods, 8 individuals
sperm whale observations: 4 pods, 8 individuals

```

The whales should be listed in alphabetical order.

All whale names should be converted to lower case.

All whale names should be converted from plural to singular - assume this can be done safely by deleting a trailing 's' if it is present.

Any extra white space should be ignored.

You can make no assumptions about possible whale names.

No mention of particular whale names can appear in your program.

As usual:

```

$ ~cs2041/bin/autotest lab06 all_whales.pl
...
$ git add all_whales.pl
$ git commit -a -m "whales & perl what could be better"

```

## Prerequisites

Write a Perl script which prints courses which can be used to meet prerequisite requirements for a UNSW course. For example:

```
$ ./prereq.pl COMP2041
COMP1917
COMP1921
$ ./prereq.pl COMP9041
COMP9021
$ ./prereq.pl COMP9242
COMP3231
COMP3891
COMP9201
COMP9283
$ ./prereq.pl HESC3641
HESC2501
```

Your script must download the UNSW handbook web pages and extract the information from them when it is run.

You should print the courses in alphabetic order.

Hints

The UNSW handbook uses separate web pages for undergraduate and postgraduate handbook and you may need to extract prerequisites from either or both.

A simple way (but not the best way) to access a web page from Perl is like this:

```
$url = "http://www.handbook.unsw.edu.au/postgraduate/courses/2015/COMP9041";
open F, "wget -q -O- $url|" or die;
while ($line = <F>) {
    print $line;
}
```

You'll have to make some assumptions about the handbook pages.

It easy in Perl to skip lines until you find one specifying prerequisites.

It easy in Perl to remove part of a line.

It easy in Perl to remove HTML tags.

```
$ ~cs2041/bin/autotest lab06 prereq.pl
...
$ git add prereq.pl
$ git commit -a -m "web scraping isn't as cool as whales"
```

## Challenge Question

Add a `-r` (for recursive ) flag to `prereq.pl` so it also also lists all courses which can be used to meet a prerequisite requirement of any course can be used to meet prerequisite requirements and so on.

Call your new program `recursive_prereq.pl` so

For example:

COMP1911  
COMP1917  
COMP1921  
COMP1927  
COMP2121  
COMP3231  
COMP3331  
COMP3891  
COMP3931  
COMP9021  
COMP9024  
COMP9032  
COMP9201  
COMP9283  
COMP9331  
ELEC1111  
ELEC1112  
ELEC2141  
ELEC2142  
MTRN2500  
MTRN3500  
TELE3018

Beware infinite loops!

Students coping well with Perl are encouraged to try picking up some Python as well.

The example Python scripts ([code/python/code\\_examples.html](#)) and links to external Python resources should help - but you may need more info - Google is your friend.

Write a Python script `digits.py` that reads from standard input and writes to standard output mapping all digit characters whose values are less than 5 into the character '`<`' and all digit characters whose values are greater than 5 into the character '`>`'. The digit character '5' should be left unchanged.

[illegible]

As usual:

```
$ ~cs2041/bin/autotest lab06 digits.py
...
$ git add digits.py
$ git commit -a -m "my first python program"
```

## Challenge Exercise: Repeated Echo in Python

Write a Perl script `echon.py` which given exactly two arguments, an integer  $n$  and a string, prints the string  $n$  times. For example:

```
$ ./echon.py 5 hello
hello
hello
hello
hello
hello
```

```
$ ./echon.py 0 nothing
```

```
$ ./echon.py 1 goodbye
goodbye
```

Your script should print an error message if it is not given appropriate arguments. For example:

```
$ ./echon.py
Usage: ./echon.py <number of lines> <string>
```

```
$ ./echon.py 1 2 3
Usage: ./echon.py <number of lines> <string>
```

```
$ ./echon.py Andrew Rocks
./echon.py: argument 1 must be a non-negative integer
```

As usual:

```
$ ~cs2041/bin/autotest lab06 echon.py
...
$ git add echon.py
$ git commit -a -m "this is a bad commit message"
```

## Challenge Exercise: Tail in Python

Write a Python script `simple_tail.py` to implement the Unix `tail` command. It does not need to implement any command line options. It will be always be given one or more files on the command line. It does not have to read from stdin or handle errors.

To assist with testing your solution, there are three small data files: `data1.txt` (`data1.txt`), `data2.txt` (`data2.txt`), and `data3.txt` (`data3.txt`). Copy these files to your current directory.

```
$ cp /home/cs2041/public_html/lab/python/simple_tail/t?.txt .
```

```
$ simple_tail.py t1.txt t2.txt t3.txt
Data 1 ... Line 2
Data 1 ... Line 3
Data 1 ... Line 4
Data 1 ... Line 5
Data 1 ... Line 6
Data 1 ... Line 7
Data 1 ... Line 8
Data 1 ... Line 9
Data 1 ... Line 10
Data 1 ... Last line
A one line file.
has
exactly
twelve
lines
with
one
word
on
each
line
```

As usual:

```
$ ~cs2041/bin/autotest lab06 simple_tail.py
...
$ git add simple_tail.py
$ git commit -a -m "my third python program"
```

## Challenge Exercise: Whales in Python

Implement Python scripts `whale.py` and `all_whales.py` which count whales as described in the first two lab questions. As usual:

```
$ ~cs2041/bin/autotest lab06 whale.py
...
$ ~cs2041/bin/autotest lab06 all_whales.py
...
$ git add whale.py all_whales.py
$ git commit -a -m "Python & whales!!!"
```

## Testing

Do your own testing in addition to the autotests.

BTW you can run all the autotests for the lab:

```
$ ~cs2041/bin/autotest lab06
```

Or just one autotest:

```
$ ~cs2041/bin/autotest lab06 whale_1
```

## Finalising

You must show your solutions to your tutor and be able to explain how they work. Once your tutor has discussed your answers with you, you should submit them using:

```
$ give cs2041 lab06 whale.pl all_whales.pl prereq.pl [recursive_prereq.pl digits.py echon.py simple_tail.py]
```

Whether you discuss your solutions with your tutor this week or next week, you must submit them before the above deadline.