Week 08 ▾ Tutorial ▾

COMP[29]041 18s2

Sample Answers ▾

1. The assignment specification doesn't fully explain the assignment - what can I do?

   A big part of the assignment is understanding what git does.
   You'll spend more time doing this than implementing subset 0.

   On line tutorials can help with this.

   You can also use the the reference implementation **2041 legit** to discover what your program is supposed to do in any situation.

2. How hard are the subsets?

   Once you understand what you have to dom subset 0 is not that hard.
   Subset 1 is hard.

   Subset 2 really hard.

   But note the marking scheme recognizes the difficulty of subsets 0 & 1.

3. What is `gitlab.cse.unsw.edu.au` and how do I want use it for ass1?

   Its a server run by CSE to host git repositories of student programs for some CSE courses.
   As assignment 1 is to implement a version control system similar to git, it is too confusing to ask you to use git for the assignment.

   Instead **give** will maintain a git repository for the assignment for you.

   Every you run give for the assignment, it will create a commit in a git repository and it will push the commit to https://gitlab.cse.unsw.edu.au/ when you can access it.

   Everytime you work on the assignment you should run give when you finish.

   This ensures you have a complete backup of all work on your program and can return to its state at any stage.

   It will also allow your tutor to check you are progressing on the lab as they can access your gitlab repository

4. What does **git init** do?
   How does this differ from **legit.pl init**

   `git init` creates an empty repository as does **legit.pl init**
   git uses the sub-directory **.git** (by default.

   legit.pl uses the sub-directory **.legit**

5. What do **git add *file*** and **legit.pl add *file*** do?

   Adds a copy of *file* to the repository's **index**.

6. What is the index in **legit** (and **git**), and where does it get stored?

   Files get added to the repositoy via the index so its somethimes called a staging area.
   It must be stored in **.legit**. exactly how you store it is up to you.

   You might create a directory **.legit/index/** and store the files there.

7. What is a commit in **legit** (and **git**), and where does it get stored?

A commit preserves the state of all files in the index.
It must be stored in **.legit**. exactly how you store it is up to you.

You might create a directory **.legit/*commit-number*** and store the files there.

8. Discuss what **./legit.pl status** prints below?

```
$ ./legit.pl init
Initialized empty legit repository in .legit
$ touch a b c d e f g h
$ ./legit.pl add a b c d e f
$ ./legit.pl commit -m 'first commit'
Commited as commit 0
$ echo hello >a
$ echo hello >b
$ echo hello >c
$ ./legit.pl add a b
$ echo world >a
$ rm d
$ ./legit.pl rm e
$ ./legit.pl add g
$ ./legit.pl status
```

```
$ ./legit.pl status
a - file changed, different changes staged for commit
b - file changed, changes staged for commit
c - file changed, changes not staged for commit
d - file deleted
e - deleted
f - same as repo
g - added to index
h - untracked
legit.pl - untracked
```

9. Write a Perl function which takes an integer argument **n** and reads the next **n** lines of input and returns them as a string.

Two sample solutions with extra code to run the function:

```perl
#!/usr/bin/perl -w

$n = shift @ARGV or die "Usage: $0 <n-lines>\n";

sub n_lines0 {
    my ($n) = @_;
    my $text = "";
    while ($n-- > 0) {
        $text .= <>;
    }
    return $text;
}

sub n_lines1 {
    my ($n) = @_;
    my $text = "";
    $text .= <> foreach (1..$n);
    return $text;
}

print n_lines1($n);
```

10. Write a Perl program which given the name of a C function searches the C source files (*.c) in the current directory for calls of the function, declarations & definitons of the function and prints a summary indicating the file and line number, in the format below. You can assume functions are defined with the type, name and paramaters on a single non-indented line. You can assume function bodies are always indented.

    You don't have to handle multi line comments. Try to avoid matching the function name in strings or single line comments. For example:

```
$ cat half.c
double half(double x) {
    return  x/2;
}
$ cat main.c
#include <stdio.h>
#include <stdlib.h>

double half(double x);

int main(int argc, char *argv[]) {
    return  half(atoi(argv[1]));
}
$ ./print_function_uses.pl half
a.c:1 function half defined
half.c:1 function half defined
main.c:4 function half declared
main.c:7 function half used
```

Perl sample solution

```perl
#!/usr/bin/perl -w

$function = $ARGV[0] or die "Usage: $0 <function-name>\n";

foreach $c_file (glob "*.c") {
    open my $cf, '<', $c_file or die "$0: can not open $c_file: $!\n";
    while ($line = <$cf>) {
        # remove single-line comments & strings (breaks if strings contain ")
        $line =~ s/\/\/.*//;
        $line =~ s/\/\*.*?\*\///;
        $line =~ s/".*?"//;
        # note use of \b (word boundary) to match function
        $line =~ /\b$function\s*\(/ or next;
        print "$c_file:$. function $function ";
        # if line is indented it should be a call to the function
        if ($line =~ /^\s/) {
            print "used\n";
        } elsif ($line =~ /;/) {
            print "declared\n";
        } else {
            print "defined\n";
        }
    }
    close $cf;
}
```

git checkout $w sample solution

```
#!/usr/bin/python
import glob, sys, re

if len(sys.argv) != 2:
    sys.stdout.write("Usage: %s <function-name>\n\n" % sys.argv[0])
    sys.exit(1)

function = sys.argv[1]

for c_file in glob.glob("*.c"):
    with open(c_file) as cf:
        # note use of \b (word boundary) to match function
        function_regex = r'\b%s\s*\(' % function
        line_number = 0
        for line in cf:
            line_number = line_number + 1
            # remove single-line comments & strings (breaks if strings contain ")
            line = re.sub(r'\/\/.*', '', line)
            line = re.sub(r'".*?"', '', line)
            if not re.search(function_regex, line):
                continue
            # if line is indented it should be a call to the function
            if re.search(r'^\s', line):
                which = "used"
            elif re.search(r';', line):
                which = "declared"
            else:
                which = "defined"
            print("%s:%d function %s %s" % (c_file, line_number, function,  which))
```

11. Write a Perl program which given a C program as input finds the definitions of single parameter functions and prints separately the function's type, name and the parameters name & type. Assume all these occur on a single non-indented line in the C source code. You can assume function bodies are always indented. Allow for white space occuring anywhere in the function header. You can assume that types in the program don't contain square or round brackets. For example:

```
$ cat a.c
double half(int *x) {
    return  *x/2.0;
}
$ ./print_function_types.pl a.c
function type='double'
function name='half'
parameter type='int *'
parameter name='x'
```

Perl sample solution

```
#!/usr/bin/perl -w

while ($line = <>) {
    $line =~ /^([a-zA-Z_].*)\((.*)\)/ or next;
    $function_start = $1;
    $parameter = $2;
    $function_type = $function_start;
    $function_type =~ s/\s*([a-zA-Z_]\w*)\s*$// or next;
    $function_name = $1;
    $parameter_type = $parameter;
    $parameter_type =~ s/\s*([a-zA-Z_]\w*)\s*$// or next;
    $parameter_name = $1;
    print "function type='$function_type'\n";
    print "function name='$function_name'\n";
    print "parameter type='$parameter_type'\n";
    print "parameter name='$parameter_name'\n";
}
```

12. Write a Perl script C_include.pl which given the name of a C source file prints the file replacing any '#include' lines with the contents of the included file, if the included file itself contains a #include line these should also be processed.

Assume the source files contain only quoted ("") include directives which contain the files's actual path name. For example:

```
$ cat f.c
#include "true.h"

int main(int argc, char *argv[]) {
    return  TRUE || FALSE;
}
$ cat true.h
#define TRUE 1
#include "false.h"
$ cat false.h
#define FALSE 0
$ ./C_include.pl f.c
#define TRUE 1
#define FALSE 0

int main(int argc, char *argv[]) {
    return  TRUE || FALSE;
}
```

Perl sample solution

```perl
#!/usr/bin/perl -w
# Given C source files interpolate #include "FILE" directives recursively.
sub include_file($);

sub include_file($) {
    my ($file) = @_;
    # this function is recursive so a local filehandle is essential
    open my $f, '<', $file or die "$0: can not open $file: $!";
    while ($line = <$f>) {
        if ($line =~ /^#\s*include\s*"([^"]*)"/) {
            include_file($1);
        } else {
            print $line;
        }
    }
    close $f;
}

foreach $file (@ARGV) {
    include_file($file);
}
```

13. Modify C_include.pl so that it handles both "" and <> directives. It should search the directories `/usr/include/`, `/usr/local/include` and `/usr/include/x86_64-linux-gnu` for include files specified in <> directives and for files specified in "" directives which do not exist locally. For example:

```
$ cat g.c
#include <stdio.h>

int main(int argc, char *argv[]) {
    printf("hello world\n");
    exit(0);
}
$ ./C_include.pl g.c|head
./C_include.pl: can not find: bits/libc-header-start.h
/* Define ISO C stdio on top of C++ iostreams.
   Copyright (C) 1991-2017 Free Software Foundation, Inc.
   This file is part of the GNU C Library.

   The GNU C Library is free software; you can redistribute it and/or
   modify it under the terms of the GNU Lesser General Public
   License as published by the Free Software Foundation; either
   version 2.1 of the License, or (at your option) any later version.

   The GNU C Library is distributed in the hope that it will be usefu
l,
```

Perl sample solution

```perl
#!/usr/bin/perl -w
# Given C source files interpolate #include "FILE" and #include <FILE>
# directives recursively.
# The recursion in this script may not terminate on stdio.h etc
# because #ifdef directive are not handled

@include_dirs = ('/usr/include/','/usr/local/include','/usr/include/x86_64-linux-gnu');

sub include_file($@);

sub include_file($@) {
    my ($file,@prefixes) = @_;
    foreach $prefix (@prefixes) {
        # this function is recursive so a local filehandle is essential
        my $path = "$prefix$file";
        next if !-r $path;
        open my $f, '<', $path or die "$0: can not open $path: $!";
        while ($line = <$f>) {
            if ($line =~ /^#\s*include\s*"(.*?)"/) {
                include_file($1, ('',@include_dirs));
            } elsif ($line =~ /^#\s*include\s*<(.*?)>/) {
                include_file($1, @include_dirs);
            } else {
                print $line;
            }
        }
        close $f;
        return;
    }
    die "$0: can not find: $file\n";
}

foreach $file (@ARGV) {
    include_file($file, (''));
}
```

Revision questions

The remaining tutorial questions are primarily intended for revision - either this week or later in session.
Your tutor may still choose to cover some of the questions time permitting.

14. Write a Perl program, `times.pl` which prints a table of multiplications.
    Your program will be given the dimension of the table and the width of the columns to be printed. For example:

```
$ ./times.pl 4 5 3
  1   1   2   3   4   5
  2   2   4   6   8  10
  3   3   6   9  12  15
  4   4   8  12  16  20
```

Sample Perl solution

```perl
#!/usr/bin/perl -w
die "Usage $0 <n> <m> <column-width>" if @ARGV != 3;
$n = $ARGV[0];
$m = $ARGV[1];
$width = $ARGV[2];
$format = "%${width}d";
foreach $x (1..$n) {
    printf $format, $x;
    foreach $y (1..$m) {
        printf "%${width}d", $x*$y;
    }
    print "\n";
}
```

Sample Python solution

```python
#!/usr/bin/python
import glob, sys, re

if len(sys.argv) != 4:
    sys.stdout.write("Usage: %s <n> <m> <column-width>\n\n" % sys.argv[0])
    sys.exit(1)

n = int(sys.argv[1])
m = int(sys.argv[2])
width = int(sys.argv[3])
format = "%%%dd" % width

for x in range(1, n + 1):
    print(format % x)
    for y in range(1, m + 1):
        sys.stdout.write(' ' + (format % (x * y)))
    print()
```

15. Write a Perl program which deletes blank lines from each of the files specified as arguments. For example, if run like this:

```
$ deblank.pl file1 file2 file3
```

your program should delete any blank lines in `file1`, `file2` and `file3`. Note that this program *changes* the files, it doesn't just write the "de-blanked" versions to standard output.

Perl sample solution

```perl
#!/usr/bin/perl -w
# delete blank lines from specified files

die "Usage: $0 <files>\n" if !@ARGV;

foreach $file (@ARGV) {
    open my $in, '<', $file or die "Can not open $file: $!";
    @lines = <$in>;  # reads entire file into array
    close $in;
    open my $out, '>', $file or die "Can not open $file: $!";
    foreach $line (@lines) {
        print $out $line if $line !~ /^\s*$/;
    }
    close $out;
}
```

Perl sample solution using -i switch

```
#!/usr/bin/perl -w -i
while (<>) {
    print if !/^\s*$/;
}
```

Perl sample solution using -i and -p switch

```
#!/usr/bin/perl -w -i -p
s/^\s*$//
```

Or from the command line:

```
$ perl -ip -e 's/^\s*$//' file1 file2 file3
```

Python sample solution - based on Perl

```
#!/usr/bin/python
# delete blank lines from specified files
# simple code which could lose data, if a write error occurs
import sys, re

for filename in sys.argv[1:]:
    with open(filename) as f:
        lines = f.readlines()
    with open(filename,'w') as f:
        for line in lines:
            if not re.match(r'^\s*$', line):
                f.write(line)
```

16. Write a Perl function `listToHTML()` that given a list of values returns a string of HTML code as an unordered list. For example

    ```
    $out = listToHTML('The', 'Quick', 'Brown', 'Fox');
    ```

    would result in `$out` having the value ...

    ```
    <ul>
    <li>The
    <li>Quick
    <li>Brown
    <li>Fox
    </ul>
    ```

    As part of an HTML page, this would display as:

    - The
    - Quick
    - Brown
    - Fox

    P.S. A Perl syntactic short cut can be used to construct the list above:

    ```
    $out = listToHTML(qw/The Quick Brown Fox/);
    ```

    Sample solution for listToHTML

    ```
    #!/usr/bin/perl -w

    sub listToHTML(@) {
        my (@list) = @_;
        return "" if !@list;
        return "<ul>\n<li>".join("\n<li>", @list)."\n</ul>\n";
    }
    print listToHTML(@ARGV);
    ```

17. Write a Perl function `hashToHTML()` that returns a string of HTML code that could be used to display a Perl associative array (hash) as an HTML table, e.g.

    ```
    # the hash table ...
    %colours = ("John"=>"blue", "Anne"=>"red", "Andrew"=>"green");
    # and the function call ...
    $out = hashToHTML(%colours);
    ```

    would result in `$out` having the value ...

```
<table border="1" cellpadding="5">
<tr><th> Key </th><th> Value </th></tr>
<tr><td> Andrew </td><td> green </td></tr>
<tr><td> Anne </td><td> red </td></tr>
<tr><td> John </td><td> blue </td></tr>
</table>
```

As part of an HTML page, this would display as:

Key	Value
Andrew	green
Anne	red
John	blue

Note that the hash should be displayed in ascending alphabetical order on key values.

This gives the function as well as some code to test it out:

```
#!/usr/bin/perl -w

sub hashToHTML {
    my (%tab) = @_;
    my $html = "";

    $html = "<table border=\"1\" cellpadding=\"5\">\n".
        "<tr><th> Key </th><th> Value </th></tr>\n";

    foreach $k (sort keys %tab) {
        $html .= "<tr><td> $k </td><td> $tab{$k} </td></tr>\n";
    }
    $html .= "</table>\n";
    return $html;
}

%h = ("David"=>"green", "Phil"=>"blue", "Andrew"=>"red", "John"=>"blue");

print hashToHTML(%h);
exit;
```

18. Write a Perl program that will read in a HTML document and output a new HTML document that contains a table with two cells (in one row). In the left cell should be a copy of the complete original HTML document inside <pre> tags so we can see the raw HTML. You will need to replace all "<" characters with the sequence "<" and all ">" characters with the sequence ">", otherwise the browser will think they are HTML tags (and we want to see the tags in the left cell). In the right cell just include the HTML body of the document, so we can see what it will look like when rendered by a browser.

    Sample solution for show_html.pl

```
#!/usr/bin/perl -w
# inspired by from www.cs.www.cs.rpi.edu/~hollingd/eiw.old/5-Perl/ex6.html

my $html_source = join "", <>;
my $modified_html = $html_source;
$modified_html =~ s/<\s*HEAD[^>]*>.*?<\s*\/HEAD[^>]*>//si;
$modified_html =~ s/<\s*\/?\s*(BODY|HTML)[^>]*>//gsi;

my ($title) = ($html_source =~ /.*<\s*TITLE[^>]*>(.*?)<\s*\/TITLE[^>]*>/si);
$title = "No title" if !defined $title;

$html_source =~ s/</\&lt;/g;
$html_source =~ s/>/\&gt;/g;

print <<eof;
<HTML>
<HEAD>
<TITLE>$title</TITLE>
</HEAD>
<BODY>
<H3 ALIGN=CENTER>HTML-VIEW of $title</H3>
<TABLE BORDER=1 BGCOLOR=WHEAT>
<TR><TD><PRE><FONT SIZE=SMALL>$html_source</FONT></PRE></TD><TD>$modified_html</TD></TR>
</TABLE>
</BODY>
</HTML>
eof
```

19. Write a Perl program that reads in data about student performance in a Prac Exam consisting of 3 exercises and computes the overall result for each student. The program takes a *single command line argument*, which is the name of a file containing space-separated text records of the form:

    ```
    studentID  exerciseID  testsPassed  numWarnings
    ```

    There will be one line in the file for each exercise submitted by a student, so a given student may have one, two or three lines of data.
    The output should be ordered by student ID and should contain a single line for each student, in the format:

    ```
    studentID  totalMark  passOrFail
    ```

    The *totalMark* value is computed as follows:

    - if an exercise passes all 5 tests, it is awarded a mark of 10 and is *correct*
    - if an exercise passes less than 5 tests, it is awarded a mark of *testsPassed/2* and is *incorrect*
    - if there are *any* warnings on an exercise, the mark is reduced by 2
    - the minimum mark for a given exercise is zero
    - the *totalMark* is the sum of the marks for the individual exercises

    The *totalMark* value should be display using the `printf` format `"%4.1f"`. A student is awarded a `PASS` if they have 2 or 3 *correct* exercises and is awarded a `FAIL` otherwise. Note that warnings do not cause an exercise to be treated as incorrect.

Sample Marks File	Corresponding Output
Command line argument: `marks1`	

`2121211 ex1 5 0`
`2121211 ex2 5 0`
`2121211 ex3 5 0`
`2233455 ex1 5 0`
`2233455 ex2 5 1`
`2233455 ex3 0 1`
`2277688 ex1 4 0`
`2277688 ex2 3 0`
`2277688 ex3 2 1`
`2277689 ex1 5 0`
`2277689 ex2 5 0`
`2277689 ex3 1 1`
	`2121211 30.0 PASS`
	`2233455 18.0 PASS`
	`2277688  3.5 FAIL`
	`2277689 20.0 PASS`

Sample Perl solution

```perl
#!/usr/local/bin/perl
#
# Prac Exam Exercise
# Author: John Shepherd (sample solution)
#

while (<>) {
    chomp;
    my ($sid,$ex,$tests,$warns) = split;
    if ($tests == 5) {
        $mark = 10;
        $ncorrect{"$sid"}++;
    }
    else {
        $mark = $tests/2.0;
    }
    $mark -= 2 if ($warns > 0);
    $mark = 0 if ($mark < 0);
    $total{$sid} += $mark;
}

foreach $sid (sort keys %total) {
    if ($ncorrect{$sid} >= 2) {
        $passfail = "PASS";
    } else {
        $passfail = "FAIL";
    }
    printf "%s %4.1f %s\n", $sid, $total{$sid}, $passfail;
}
```

Sample Python solution

```python
#!/usr/bin/python
import fileinput, re, sys, collections
ncorrect = collections.defaultdict(int)
total = collections.defaultdict(float)
for line in fileinput.input():
    (sid,ex,tests,warns) = line.split()
    if tests == '5':
        mark = 10
        ncorrect[sid] += 1
    else:
        mark = int(tests)/2.0
    if int(warns) > 0:
        mark = max(0, mark - 2)
    total[sid] += mark
for sid in sorted(total.keys()):
    if ncorrect[sid] >= 2:
        passfail = "PASS"
    else:
        passfail = "FAIL"
    print("%s %4.1f %s" % (sid, total[sid], passfail))
```

20. What does this Perl print and why?

```perl
@a = (1..5);
@b = grep { $_ = $_ - 3; $_ > 0} @a;
print "@a\n";
print "@b\n";
```

It prints:

```
-2 -1 0 1 2
1 2
```

The `grep` function aliases `$_` to each list element in turn an executes the code in the block. It returns a list ofthe element for which the last expressione evaluated is true.

`{ $_ = $_ - 3 }` subtracts 3 from each element in `@a`. The `$_ > 0` expression selects positive elements.

21. What does this Perl print?

```
@vec = map { $_ ** 2 } (1,2,3,4,5);
print "@vec\n";
```

It prints:

```
1 4 9 16 25
```

The `map` function applies the code in the block `{ $_ ** 2 }` to each element in the list, and returns a list containing the tranformed values. The `**` operator does exponentiation; and `$_` refers to the "current" element in the list.