CS3157: Advanced Programming

Lecture #4

Jan 30

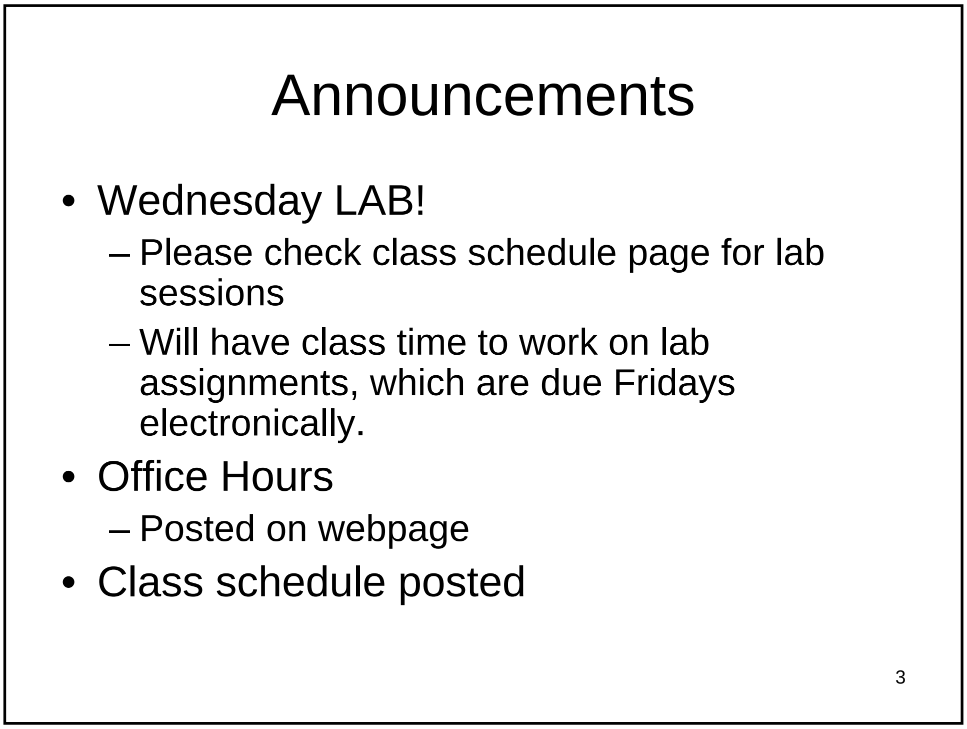
Shlomo Hershkop

shlomo@cs.columbia.edu

Outline

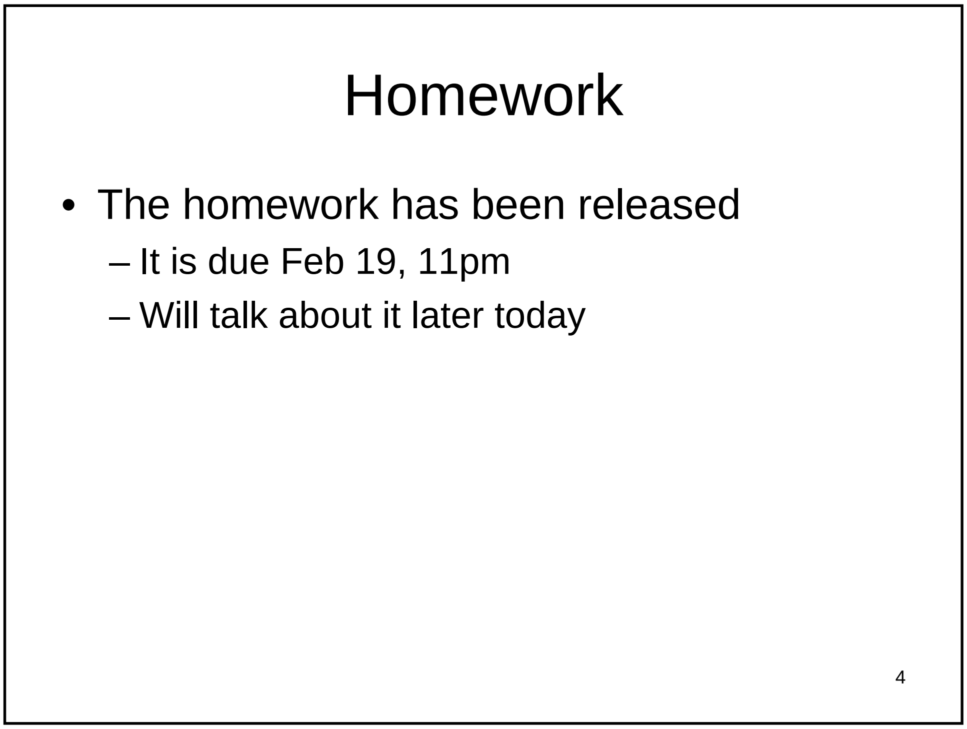
- Feedback
- Homework
- More file handling and reg exp
- CGI
- HTML
- CGI & Perl
- Perl Debugger

- Reading:
  - Regular expressions
  - File handling



Announcements

- Wednesday LAB!
  - Please check class schedule page for lab sessions
  - Will have class time to work on lab assignments, which are due Fridays electronically.
- Office Hours
  - Posted on webpage
- Class schedule posted



Homework

- The homework has been released
  - It is due Feb 19, 11pm
  - Will talk about it later today

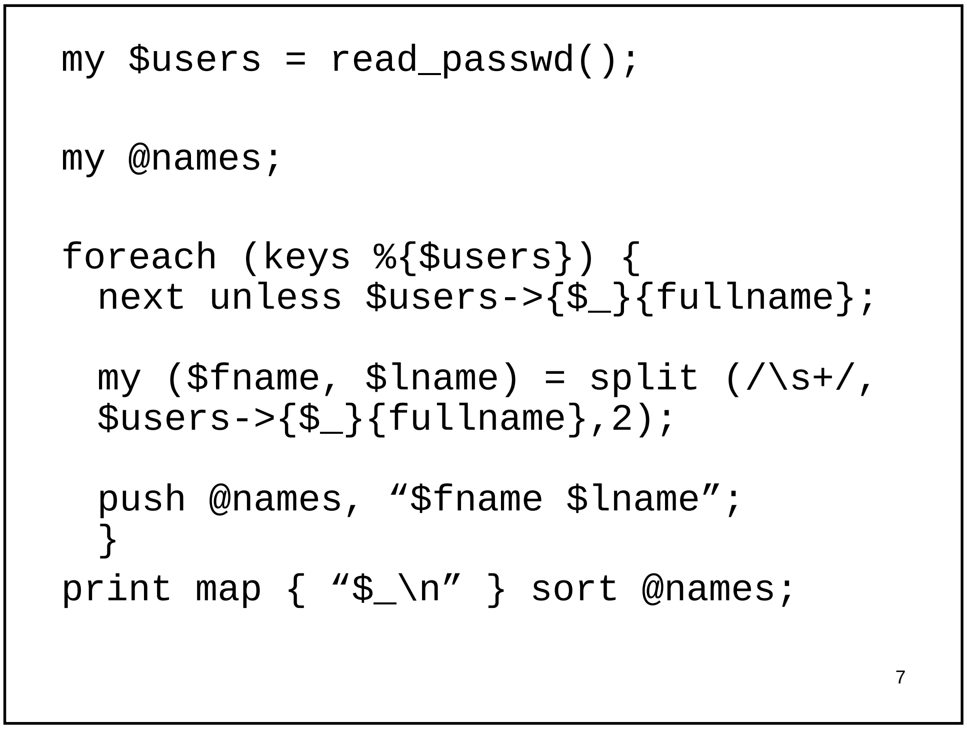
More code examples

- We want to process the /etc/password file

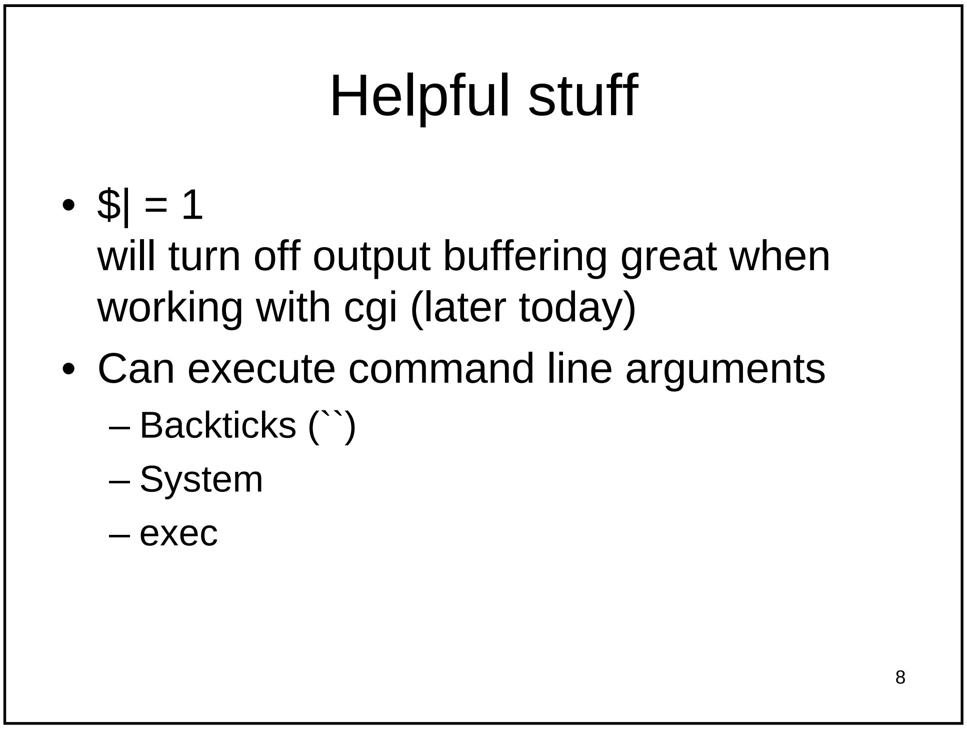
- Looks like:

```
pcap:x:77:77:ARPWATCH User:/var/arpwatch:/sbin/nologin
ident:x:98:98:pident user:/:/bin/false
nobody:x:99:99:Nobody:/:/sbin/nologin
xfs:x:405:405:X Font Server:/etc/X11/fs:/bin/false
mysql:x:6730:1101:mysql server:/var/lib/mysql:/bin/bash
```

```
sub read_passwd {
my %users;
my @fields = qw/name pword uid gid fullname
   home shell/;

while(<STDIN>) {
   chomp;
   my %rec;

   @rec ={@fields} = split(/:/);

   $users{$rec{name}} = \%rec;
   }
   return \%users;
}
```



```
my $users = read_passwd();

my @names;

foreach (keys %{$users}) {
  next unless $users->{$_}{fullname};

  my ($fname, $lname) = split (/\s+/,
  $users->{$_}{fullname},2);

  push @names, “$fname $lname”;
  }
print map { “$_\n” } sort @names;
```



Helpful stuff

- $| = 1
  will turn off output buffering great when working with cgi (later today)
- Can execute command line arguments
  - Backticks (``)
  - System
  - exec

MD5 Sum

- MD5 – uses a 128 bit hash value
- Designed in 1991
- Known problems with collision attacks
- http://www.ietf.org/rfc/rfc1321.txt
- http://en.wikipedia.org/wiki/MD5

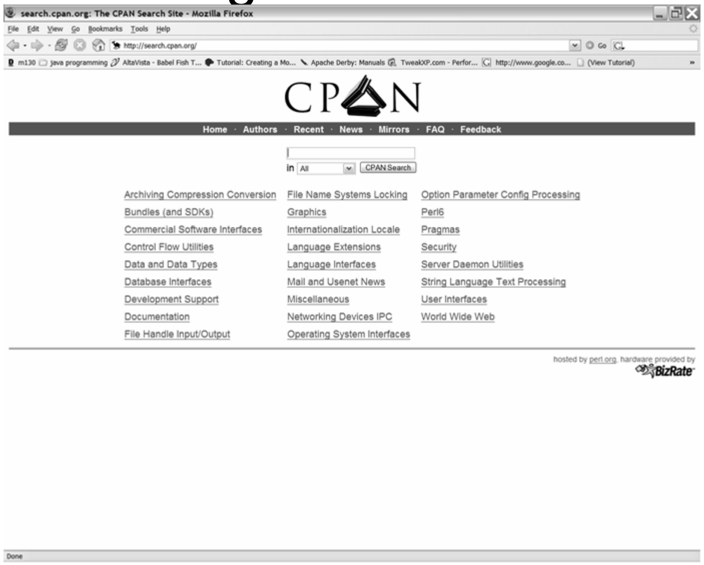
Bottom line

- Still in very wide use
- Allows authentication of files given a file and signature
- Visually authentication against tampering
- What obvious weakness??

Md5 of a file

- Can execute md5sum within perl
- Can use perl defined methods
  - – Write yourself
  - – Find someone else's ☺

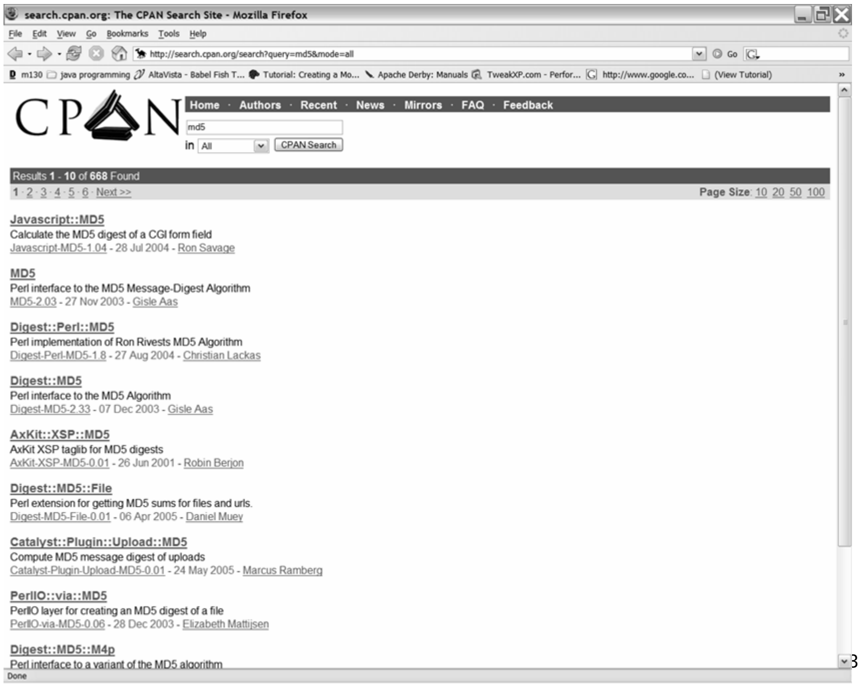
Using Perl Libraries


search.cpan.org: The CPAN Search Site - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://search.cpan.org/
Go
m130 java programming AltaVista - Babel Fish T... Tutorial: Creating a Mo... Apache Derby: Manuals TweakXP.com - Perfor... http://www.google.co... (View Tutorial)
CPAN
Home · Authors · Recent · News · Mirrors · FAQ · Feedback
in All
CPAN Search
hosted by perl.org, hardware provided by BizRate
Done

Archiving Compression Conversion
Bundles (and SDKs)
Commercial Software Interfaces
Control Flow Utilities
Data and Data Types
Database Interfaces
Development Support
Documentation
File Handle Input/Output

File Name Systems Locking
Graphics
Internationalization Locale
Language Extensions
Language Interfaces
Mail and Usenet News
Miscellaneous
Networking Devices IPC
Operating System Interfaces

Option Parameter Config Processing
Perl6
Pragmas
Security
Server Daemon Utilities
String Language Text Processing
User Interfaces
World Wide Web


search.cpan.org: The CPAN Search Site - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://search.cpan.org/search?query=md5&mode=all
Go
m130 java programming AltaVista - Babel Fish T... Tutorial: Creating a Mo... Apache Derby: Manuals TweakXP.com - Perfor... http://www.google.co... (View Tutorial)
CPAN
Home · Authors · Recent · News · Mirrors · FAQ · Feedback
md5
in All
CPAN Search
Done

Results **1** - **10** of **668** Found

1 · 2 · 3 · 4 · 5 · 6 · Next >>

Page Size: 10 20 50 100

Javascript::MD5
Calculate the MD5 digest of a CGI form field
Javascript-MD5-1.04 - 28 Jul 2004 - Ron Savage

MD5
Perl interface to the MD5 Message-Digest Algorithm
MD5-2.03 - 27 Nov 2003 - Gisle Aas

Digest::Perl::MD5
Perl implementation of Ron Rivests MD5 Algorithm
Digest-Perl-MD5-1.8 - 27 Aug 2004 - Christian Lackas

Digest::MD5
Perl interface to the MD5 Algorithm
Digest-MD5-2.33 - 07 Dec 2003 - Gisle Aas

AxKit::XSP::MD5
AxKit XSP taglib for MD5 digests
AxKit-XSP-MD5-0.01 - 26 Jun 2001 - Robin Berjon

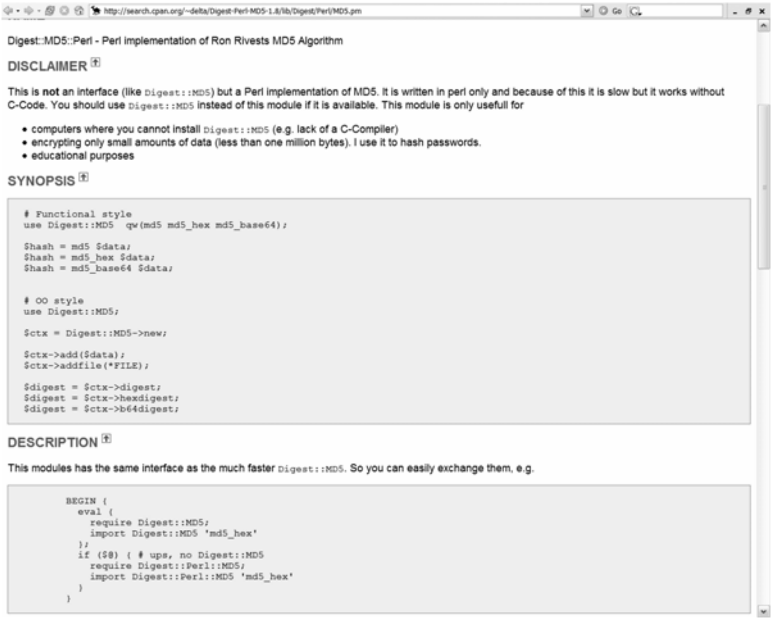
Digest::MD5::File
Perl extension for getting MD5 sums for files and urls.
Digest-MD5-File-0.01 - 06 Apr 2005 - Daniel Muey

Catalyst::Plugin::Upload::MD5
Compute MD5 message digest of uploads
Catalyst-Plugin-Upload-MD5-0.01 - 24 May 2005 - Marcus Ramberg

PerlIO::via::MD5
PerlIO layer for creating an MD5 digest of a file
PerlIO-via-MD5-0.06 - 28 Dec 2003 - Elizabeth Mattijsen

Digest::MD5::M4p
Perl interface to a variant of the MD5 algorithm

3


http://search.cpan.org/~delta/Digest-Perl-MD5-1.8/lib/Digest/Perl/MD5.pm
Go

Digest::MD5::Perl - Perl implementation of Ron Rivests MD5 Algorithm

DISCLAIMER

This is **not** an interface (like `Digest::MD5`) but a Perl implementation of MD5. It is written in perl only and because of this it is slow but it works without C-Code. You should use `Digest::MD5` instead of this module if it is available. This module is only usefull for

- computers where you cannot install `Digest::MD5` (e.g. lack of a C-Compiler)
- encrypting only small amounts of data (less than one million bytes). I use it to hash passwords.
- educational purposes

SYNOPSIS

```
# Functional style
use Digest::MD5  qw(md5 md5_hex md5_base64);

$hash = md5 $data;
$hash = md5_hex $data;
$hash = md5_base64 $data;


# OO style
use Digest::MD5;

$ctx = Digest::MD5->new;

$ctx->add($data);
$ctx->addfile(*FILE);

$digest = $ctx->digest;
$digest = $ctx->hexdigest;
$digest = $ctx->b64digest;
```

DESCRIPTION

This modules has the same interface as the much faster `Digest::MD5`. So you can easily exchange them, e.g.

```
    BEGIN {
      eval {
        require Digest::MD5;
        import Digest::MD5 'md5_hex'
      };
      if ($@) { # ups, no Digest::MD5
        require Digest::Perl::MD5;
        import Digest::Perl::MD5 'md5_hex'
      }
    }
```

14

Digests

- The 128-bit (16-byte) MD5 hashes (also termed message digests) are typically represented as 32-digit hexadecimal numbers.
- Even small change can result in a totally different hash digest

Digests II

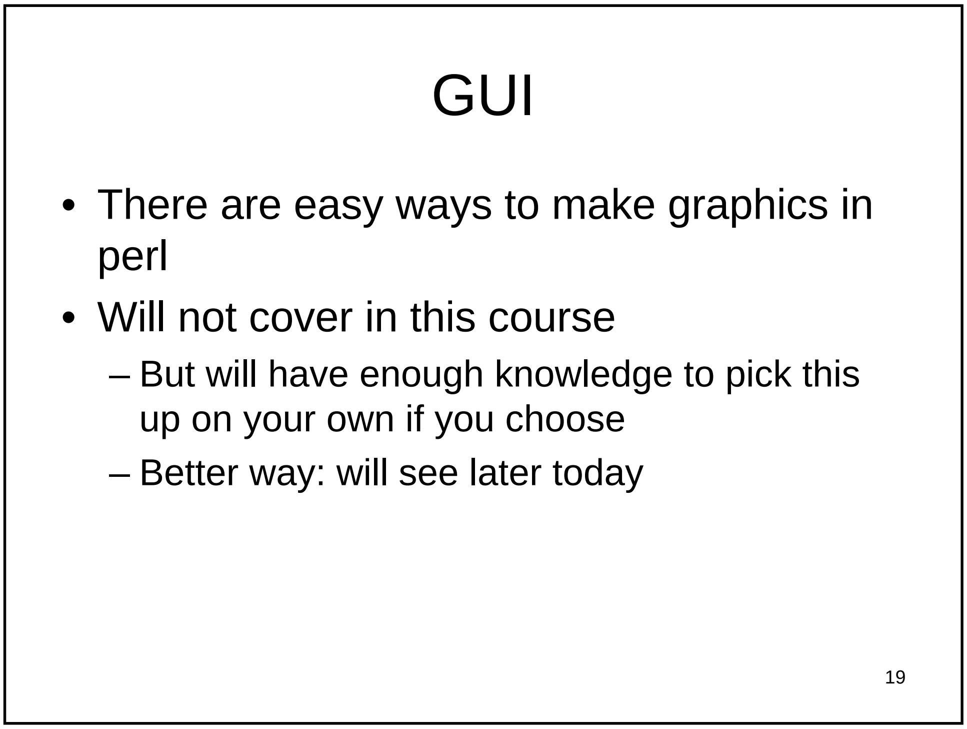
- MD5("The quick brown fox jumps over the lazy dog") =
  - 9e107d9d372bb6826bd81d3542a419d6
- MD5("The quick brown fox jumps over the lazy cog") =
  - 1055d3e698d289f2af8663725127bd4b
- MD5(“”)
  - d41d8cd98f00b204e9800998ecf8427e

Recursive directory crawling

- Sample1.pl

File::Find

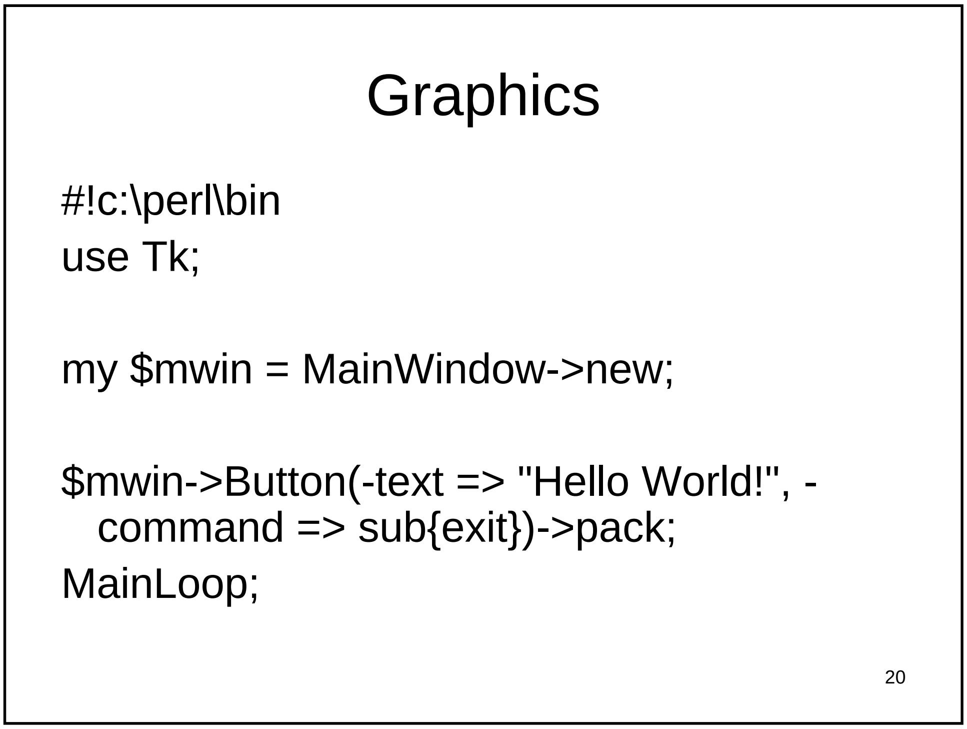
```
use File::Find;

  $dir = “c:/example”;


find(\&exam1,$dir);

  sub exam1{
     print “File: $_ and path is
  $File::Find::name\n”;
}
```



GUI

- There are easy ways to make graphics in perl
- Will not cover in this course
  - But will have enough knowledge to pick this up on your own if you choose
  - Better way: will see later today

19



Graphics

```
#!c:\perl\bin
use Tk;

my $mwin = MainWindow->new;

$mwin->Button(-text => "Hello World!", -
   command => sub{exit})->pack;
MainLoop;
```

20

Graphics

- Will not cover in depth
- Good to know about
- Might need to one day debug someone else’s code (GASP!)

Computer Security

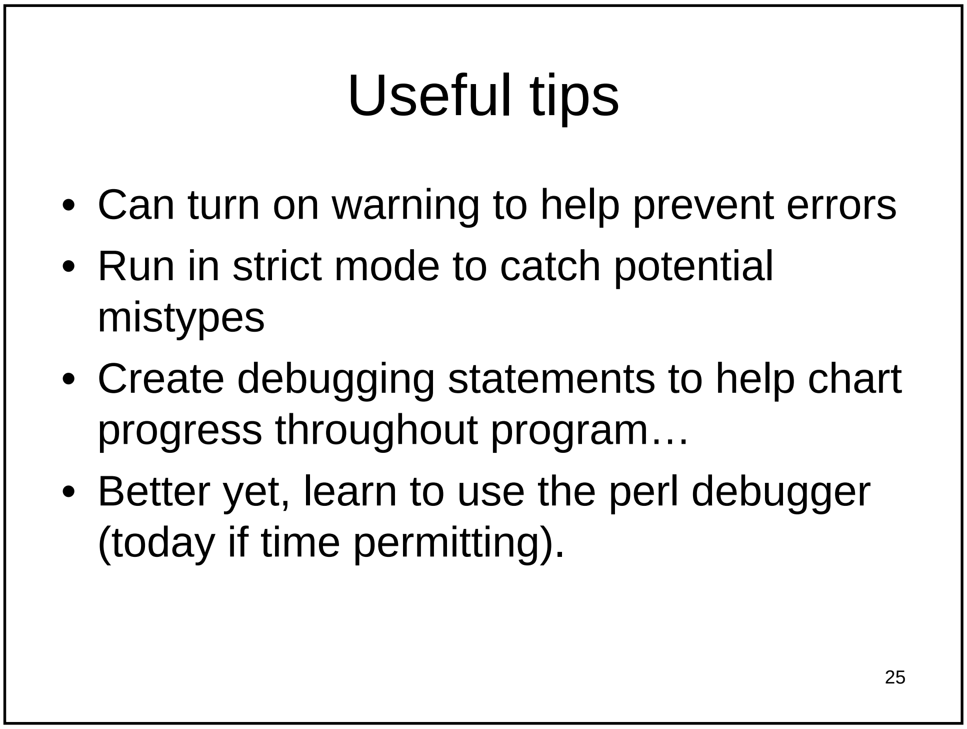
- System and theory of ensuring the confidentiality, integrity, availability, and control of electronic information and systems.
  – Network
  – Host
  – Data

For host based security

- Want to ensure permission system
  - X should only be allowed to do A, B, and C
- Want to ensure accountability
  - If Y does something not allowed, should be noted
- Want to be able to track
  - If something has been tampered with, how can we locate it
  - Both preventative and reactionary

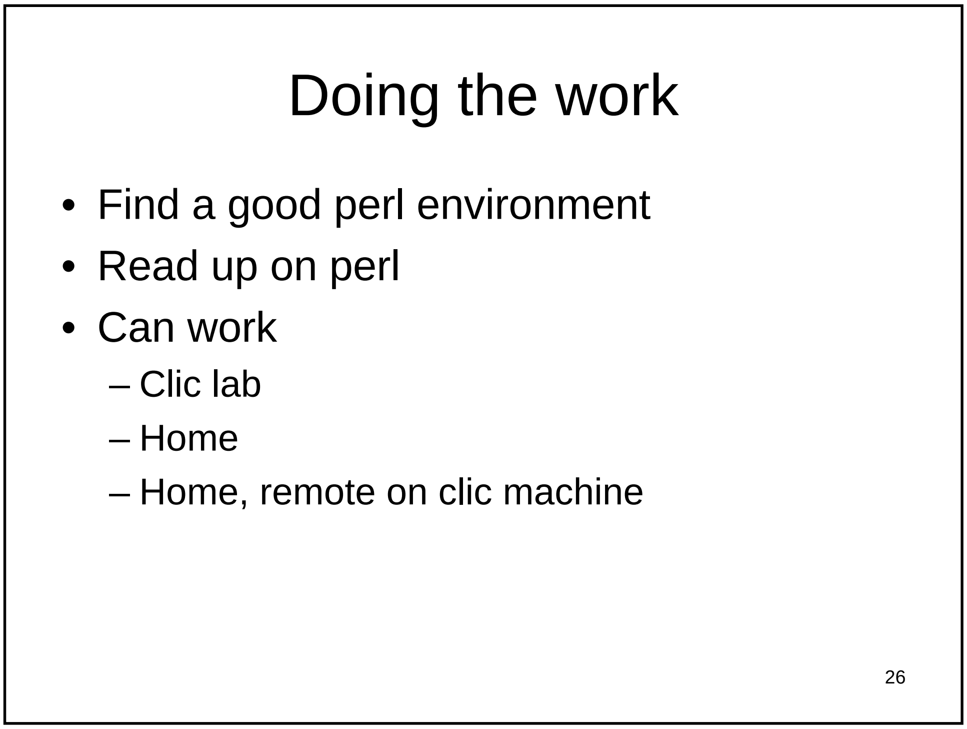
Homework Project

- Assuming you are a system administrator or just paranoid
- Take chronological snapshots of your system to compare and find changes
  - Many changes by system
  - Many changes by valid user
  - Might locate malicious user/system changes



Useful tips

- Can turn on warning to help prevent errors
- Run in strict mode to catch potential mistypes
- Create debugging statements to help chart progress throughout program…
- Better yet, learn to use the perl debugger (today if time permitting).



Doing the work

- Find a good perl environment
- Read up on perl
- Can work
  - Clic lab
  - Home
  - Home, remote on clic machine

TOOLS: VNC

- www.realvnc.com
- Start server on a clic machine:
  – vncserver

  – Run client on your side

  – demo

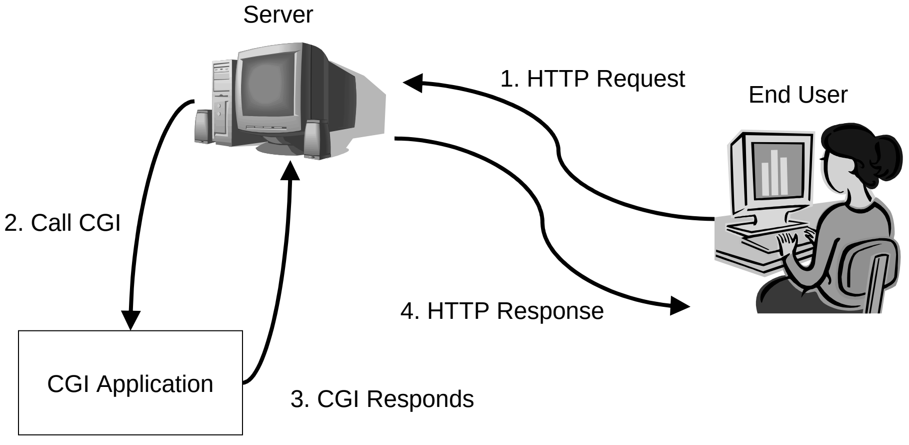
www

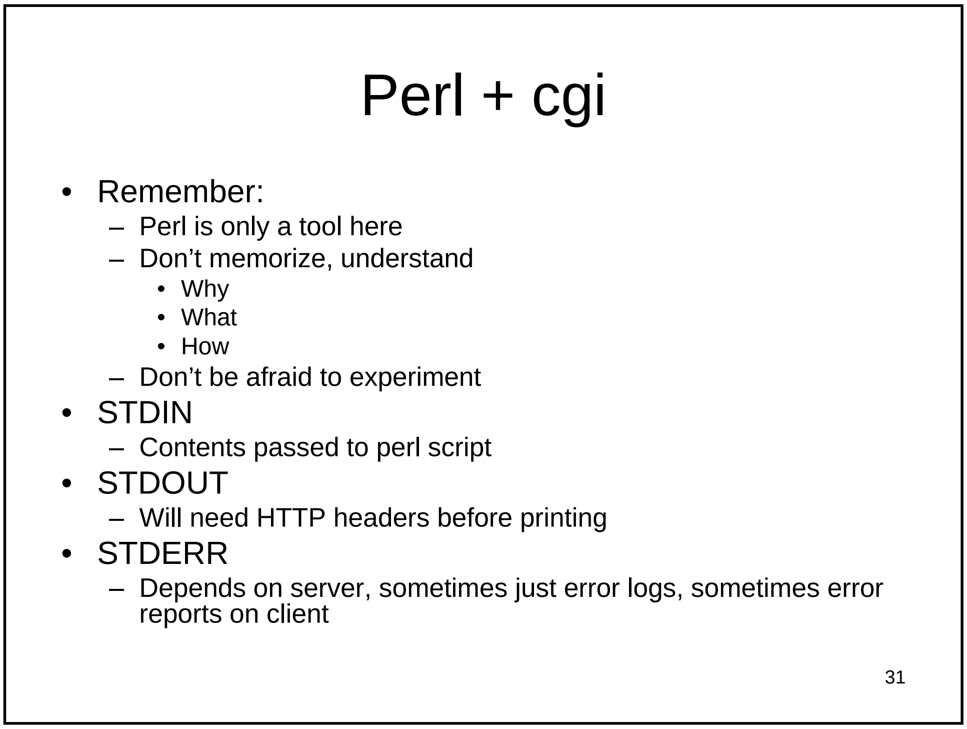
- Driven by http
- Technical overview
  – Servers serve http request
  – Clients browsers issue requests

Boring vs. Exciting

- Typical
  - Request is served from a file formatted in html
  - Static file of what we would like to render on a web client.
  - Example:
    - Class syllabus
- What is we could tailor each users web experience to what they want.
  - Design of protocol to handle this

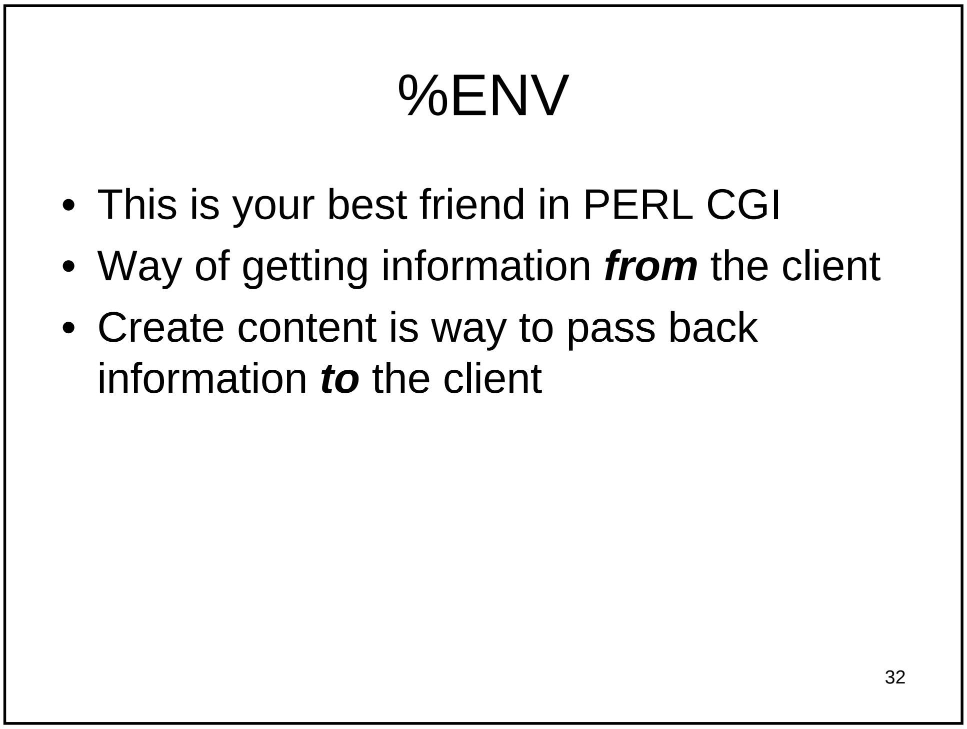
How does CGI work:





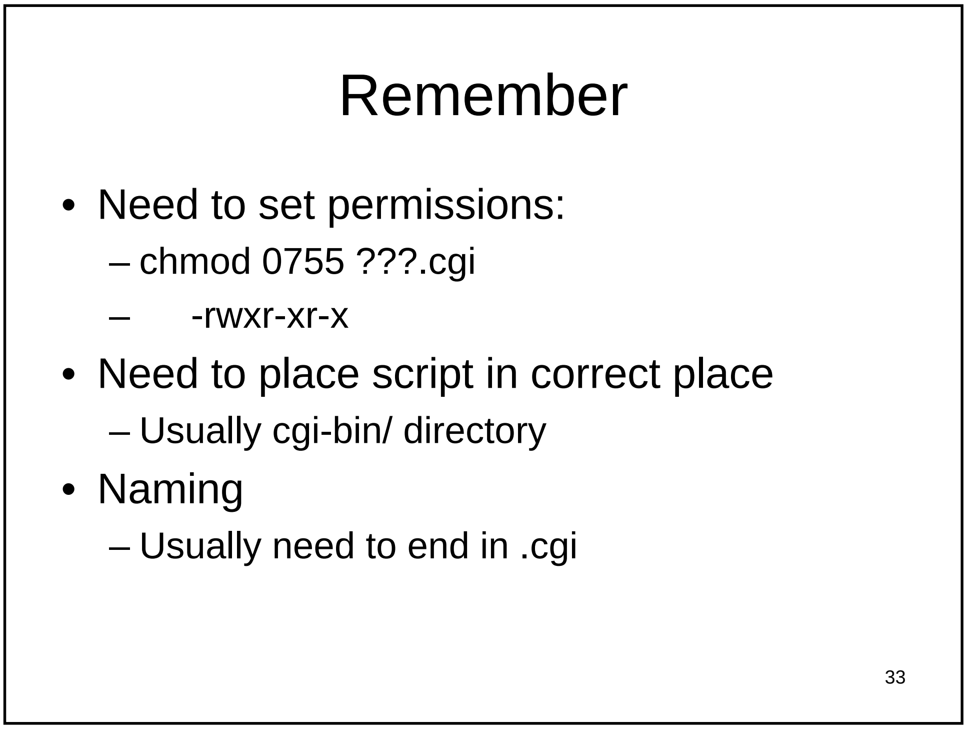
Perl + cgi

- Remember:
  - Perl is only a tool here
  - Don’t memorize, understand
    - Why
    - What
    - How
  - Don’t be afraid to experiment
- STDIN
  - Contents passed to perl script
- STDOUT
  - Will need HTTP headers before printing
- STDERR
  - Depends on server, sometimes just error logs, sometimes error reports on client



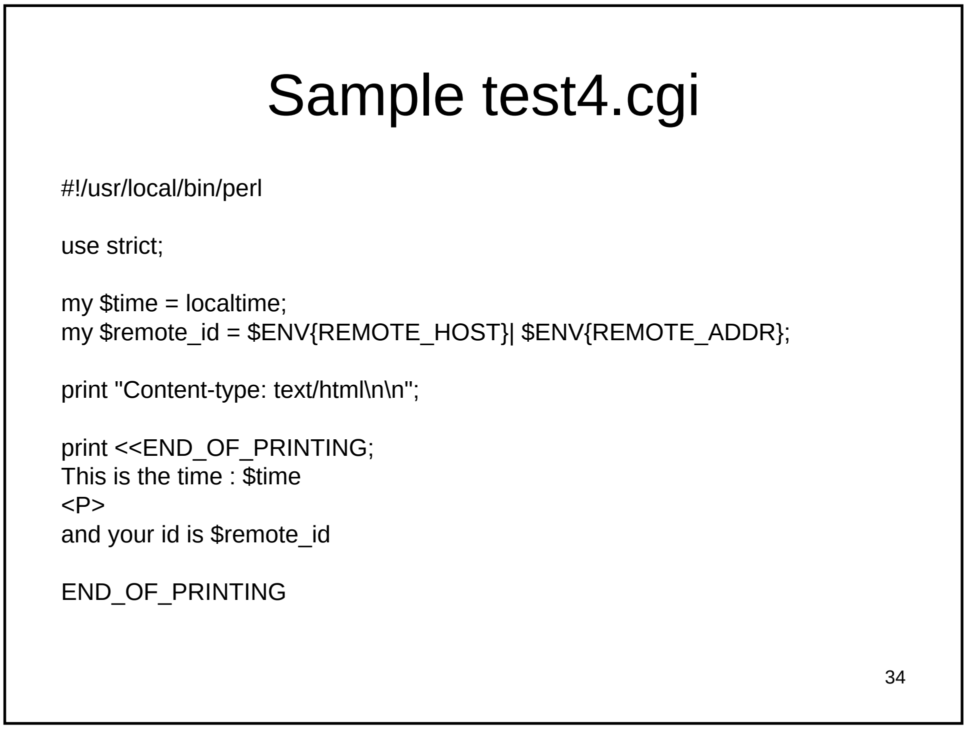
%ENV

- This is your best friend in PERL CGI
- Way of getting information ***from*** the client
- Create content is way to pass back information ***to*** the client



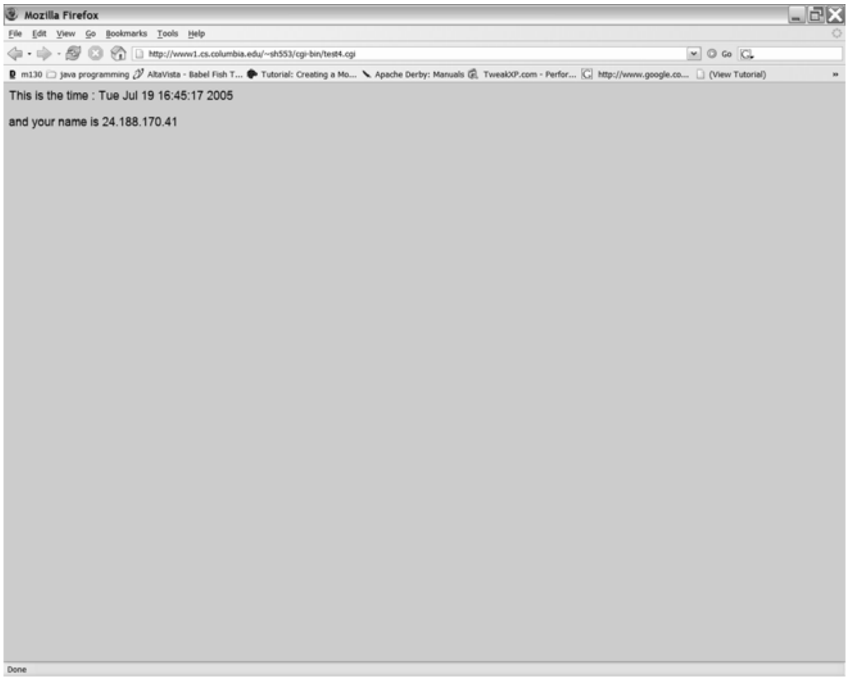
Remember

- Need to set permissions:
  - chmod 0755 ???.cgi
  - -rwxr-xr-x
- Need to place script in correct place
  - Usually cgi-bin/ directory
- Naming
  - Usually need to end in .cgi



Sample test4.cgi

```
#!/usr/local/bin/perl

use strict;

my $time = localtime;
my $remote_id = $ENV{REMOTE_HOST}| $ENV{REMOTE_ADDR};

print "Content-type: text/html\n\n";

print <<END_OF_PRINTING;
This is the time : $time
<P>
and your id is $remote_id

END_OF_PRINTING
```


Mozilla Firefox
File Edit View Go Bookmarks Tools Help
Go
m130 java programming AltaVista - Babel Fish T... Tutorial: Creating a Mo... Apache Derby: Manuals TweakXP.com - Perfor... http://www.google.co... (View Tutorial)
Done

http://www1.cs.columbia.edu/~sh553/cgi-bin/test4.cgi

This is the time : Tue Jul 19 16:45:17 2005

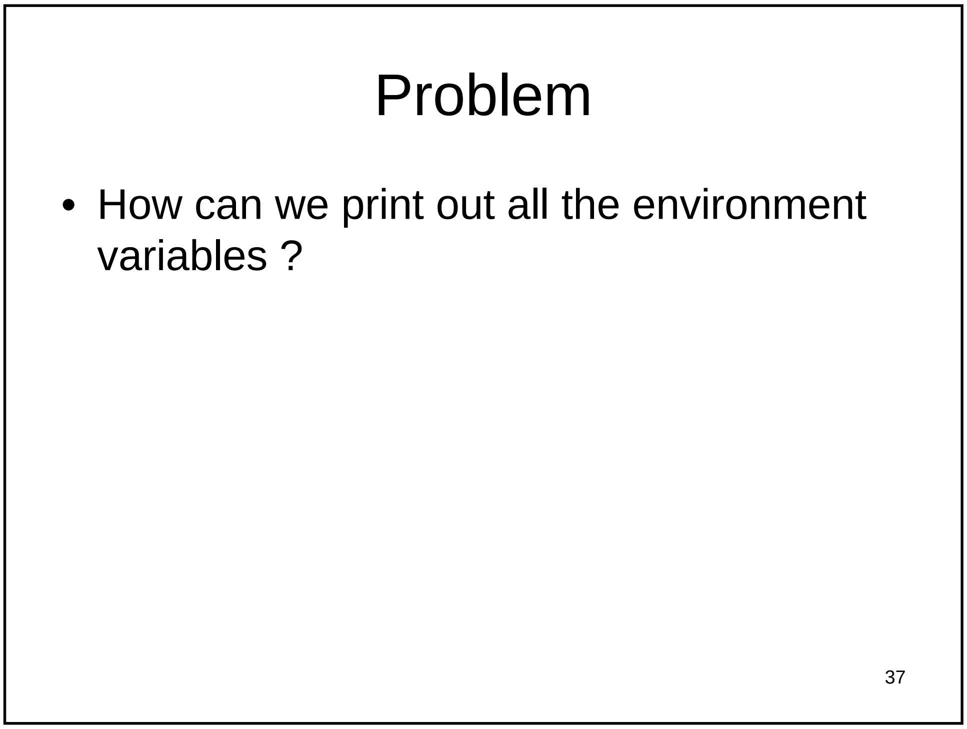
and your name is 24.188.170.41

35

Some CGI Environmental Variables

- CONTENT_LENGTH
  - Length of data passed to cgi
- CONTENT_TYPE
- QUERY_STRING
- REMOTE_ADDR
  - Ip address of client
- REQUEST_METHOD
- SCRIPT_NAME
- SERVER_PORT
- SERVER_NAME
- SERVER_SOFTWARE
- HTTP_FROM
- HTTP_USER_AGENT
- HTTP_REFERER
- HTTP_ACCEPT

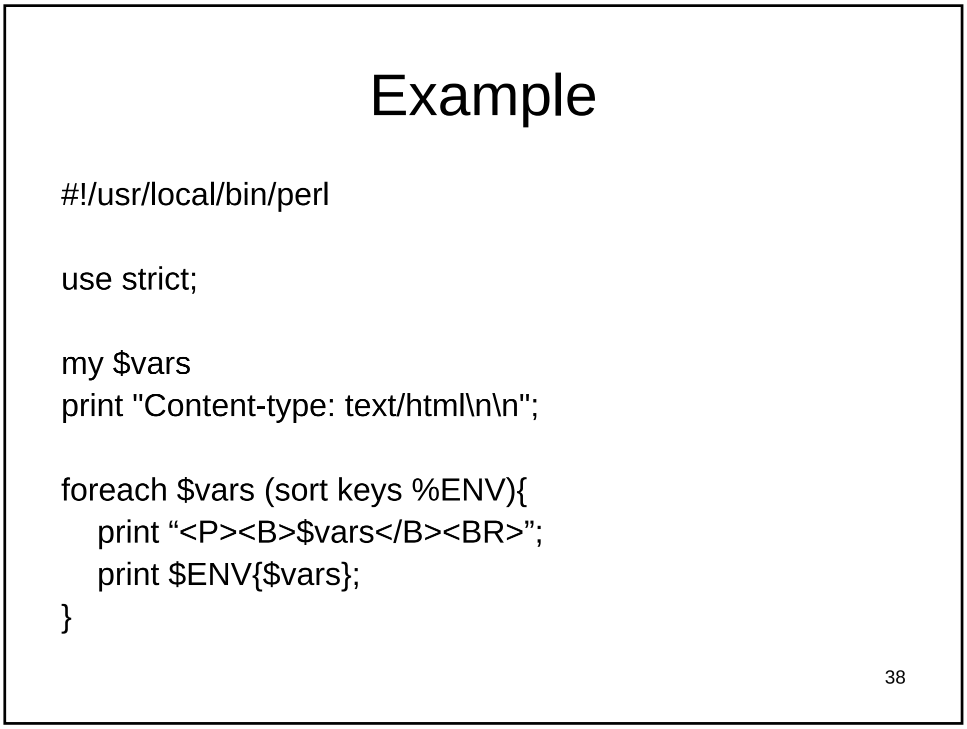
36



Problem

- How can we print out all the environment variables ?

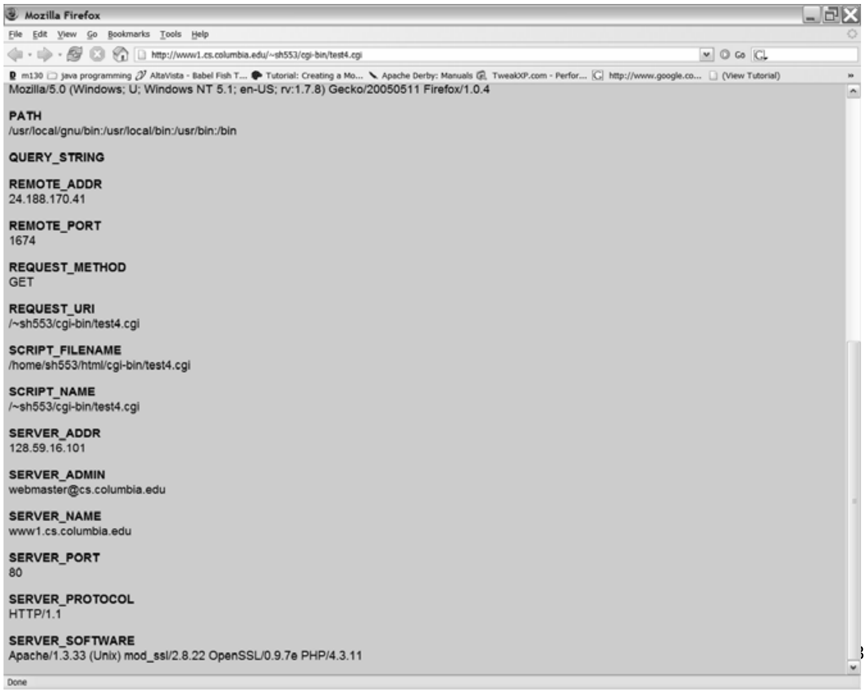
37



Example

```
#!/usr/local/bin/perl

use strict;

my $vars
print "Content-type: text/html\n\n";

foreach $vars (sort keys %ENV){
    print “<P><B>$vars</B><BR>”;
    print $ENV{$vars};
}
```

38


Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://www1.cs.columbia.edu/~sh553/cgi-bin/test4.cgi
Go
m130 java programming AltaVista - Babel Fish T... Tutorial: Creating a Mo... Apache Derby: Manuals TweakXP.com - Perfor... http://www.google.co... (View Tutorial)
Done

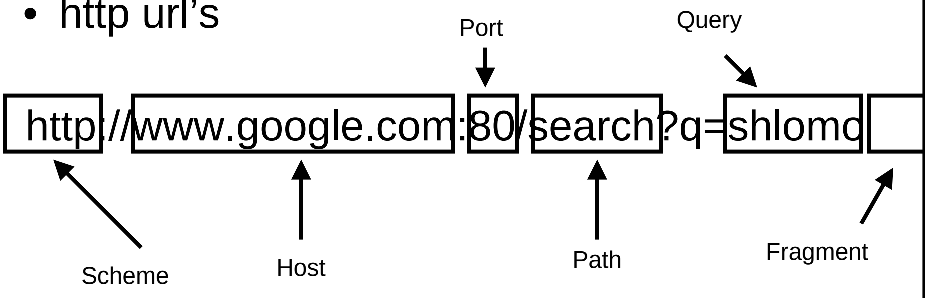
```
Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.7.8) Gecko/20050511 Firefox/1.0.4

PATH
/usr/local/gnu/bin:/usr/local/bin:/usr/bin:/bin

QUERY_STRING

REMOTE_ADDR
24.188.170.41

REMOTE_PORT
1674

REQUEST_METHOD
GET

REQUEST_URI
/~sh553/cgi-bin/test4.cgi

SCRIPT_FILENAME
/home/sh553/html/cgi-bin/test4.cgi

SCRIPT_NAME
/~sh553/cgi-bin/test4.cgi

SERVER_ADDR
128.59.16.101

SERVER_ADMIN
webmaster@cs.columbia.edu

SERVER_NAME
www1.cs.columbia.edu

SERVER_PORT
80

SERVER_PROTOCOL
HTTP/1.1

SERVER_SOFTWARE
Apache/1.3.33 (Unix) mod_ssl/2.8.22 OpenSSL/0.9.7e PHP/4.3.11
```

HTML

- Hyper Text Markup Language
- Standard by w3: http://www.w3.org/MarkUp/
- Way of standardizing format of documents so that users can share information between different systems seamlessly
- Evolving to XHTML format

HTML

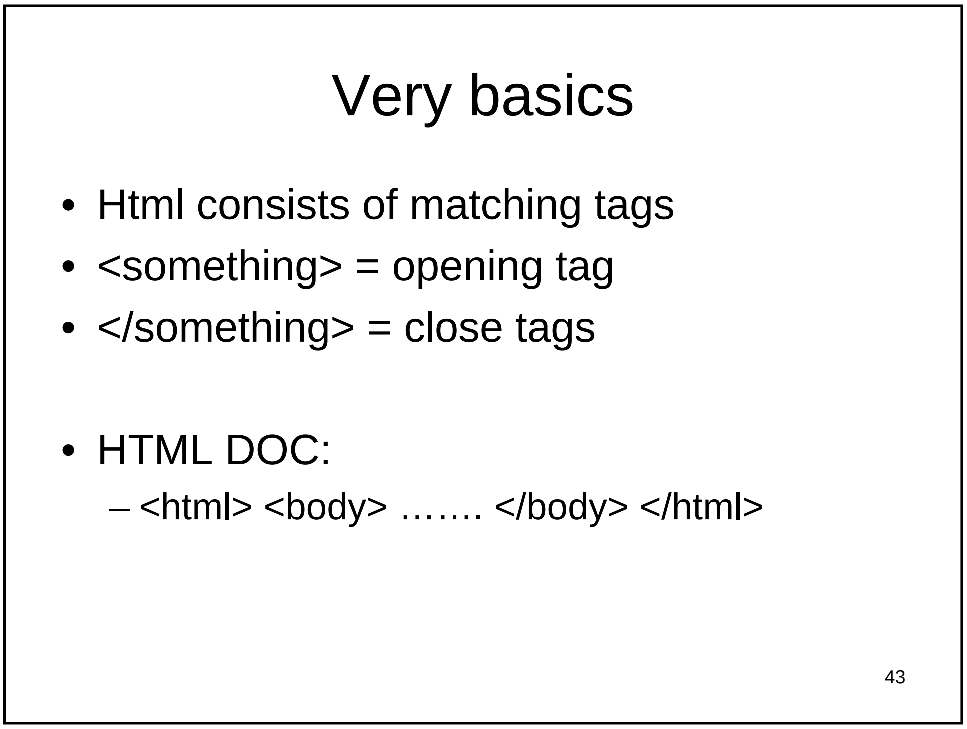
- Hypertext Transfer Protocol
- Language used between web servers and web clients
- http url's



http://www.google.com:80/search?q=shlomo

Google.com

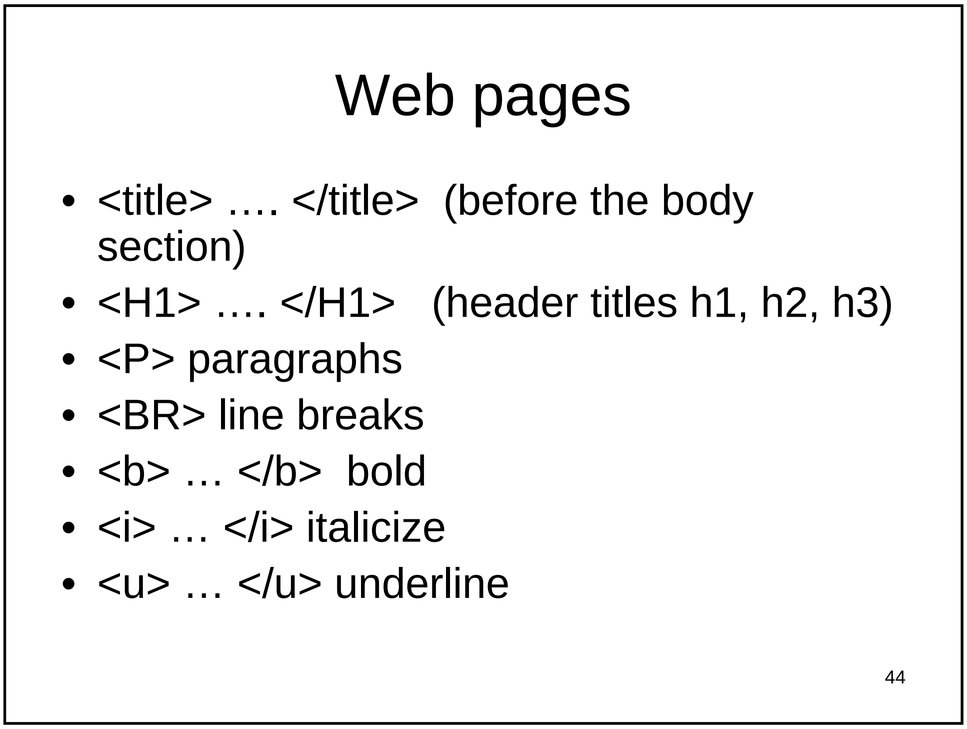
- http://www.google.com/search?q=shlomo



Very basics

- Html consists of matching tags
- <something> = opening tag
- </something> = close tags

- HTML DOC:
  – <html> <body> ……. </body> </html>



Web pages

- <title> …. </title> (before the body section)
- <H1> …. </H1> (header titles h1, h2, h3)
- <P> paragraphs
- <BR> line breaks
- <b> … </b> bold
- <i> … </i> italicize
- <u> … </u> underline

More basics

- <img src =“…..” width=“X” height=“Y”>
- <a href=“www.cnn.com”> something </a>
- <a name=“Anchor1”>
  - – Can be referred to by page.html#Anchor1
- <hr>   line
- <hr width=50%> half line

Lists

- Unordered list

<ul> <li> </li> ……</ul>

- Ordered list

<ol> <li> </li> ….. </ol>

- Nested lists
  - – Lists themselves can be nested within another

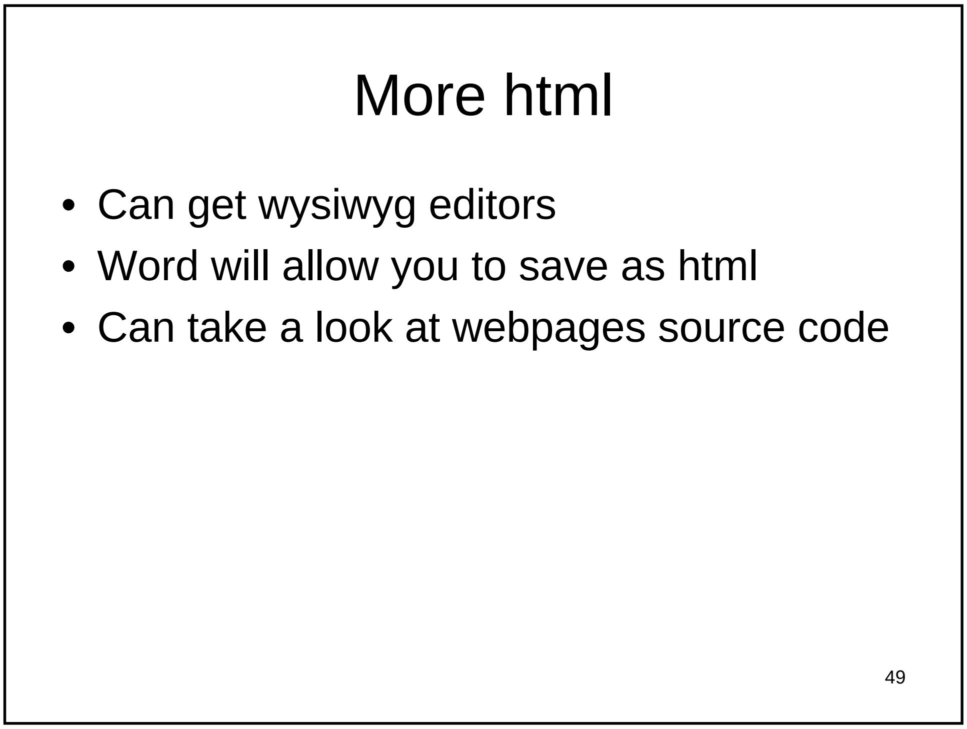
Tables

-
  ```
  <table>
  <tr>
  <td>Hello</td>
  <td>World </td>
  </tr>
  </table>
  ```

Hello	World

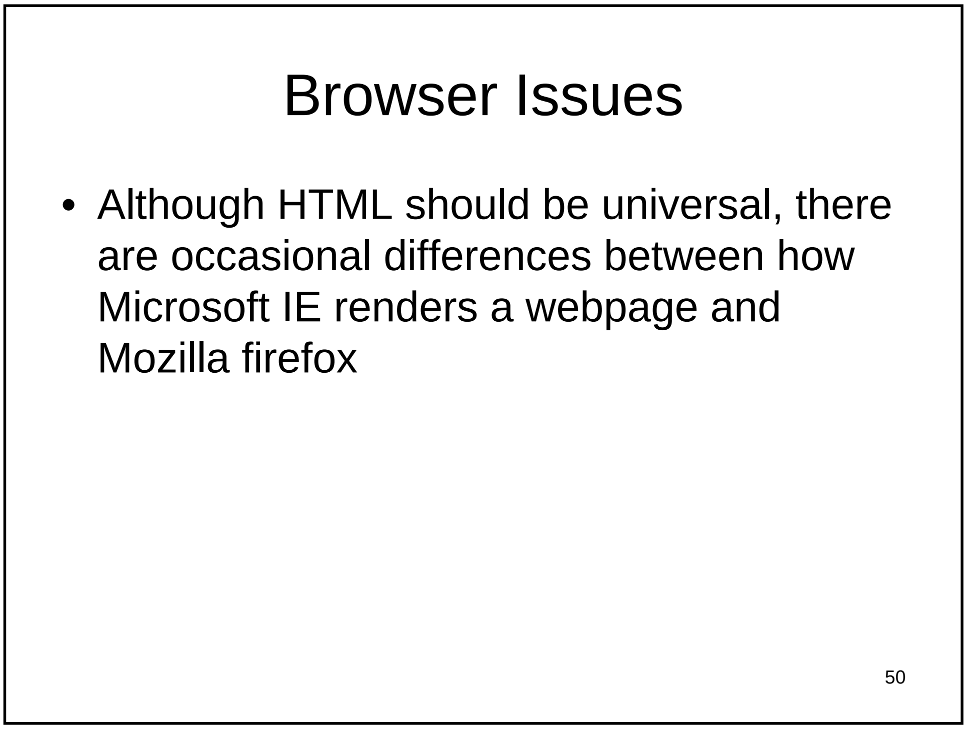
comments

```
<!--

   anything you do

   -->
```



More html

- Can get wysiwyg editors
- Word will allow you to save as html
- Can take a look at webpages source code



Browser Issues

- Although HTML should be universal, there are occasional differences between how Microsoft IE renders a webpage and Mozilla firefox

Perl Debugging

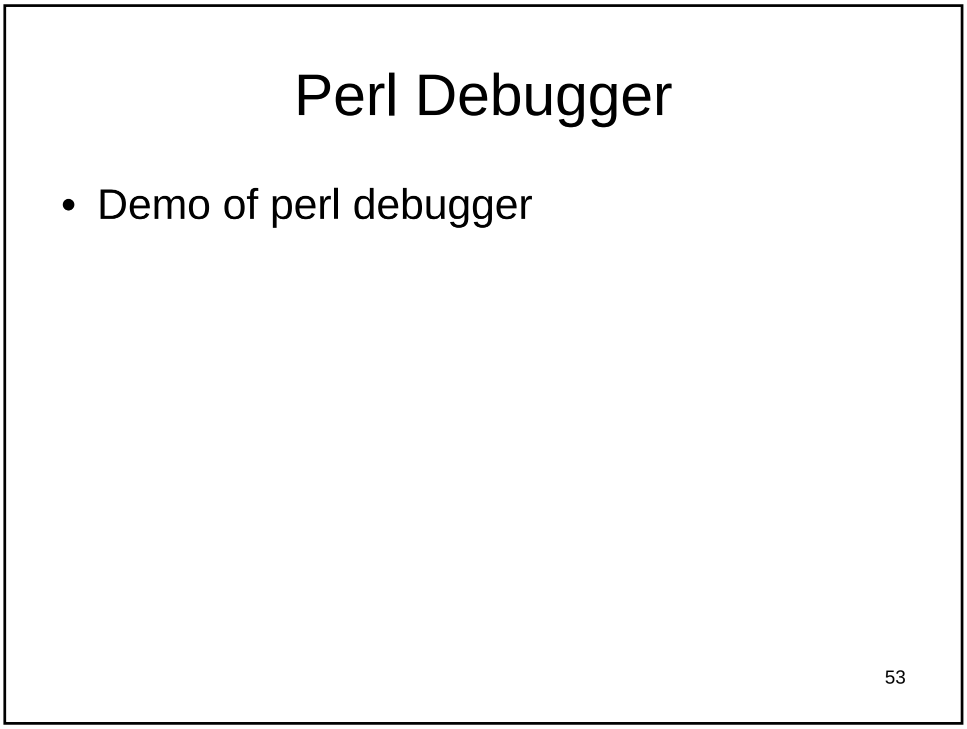
- Command line debugger can be started with the -d command argument

perl –d something.pl

- h = help
- x = examine something
- Any perl command is read in, and saved
- s = single step evaluation
- n = jump over subroutine
- v [num] = window of commands we are in
- l x y = list lines x to y

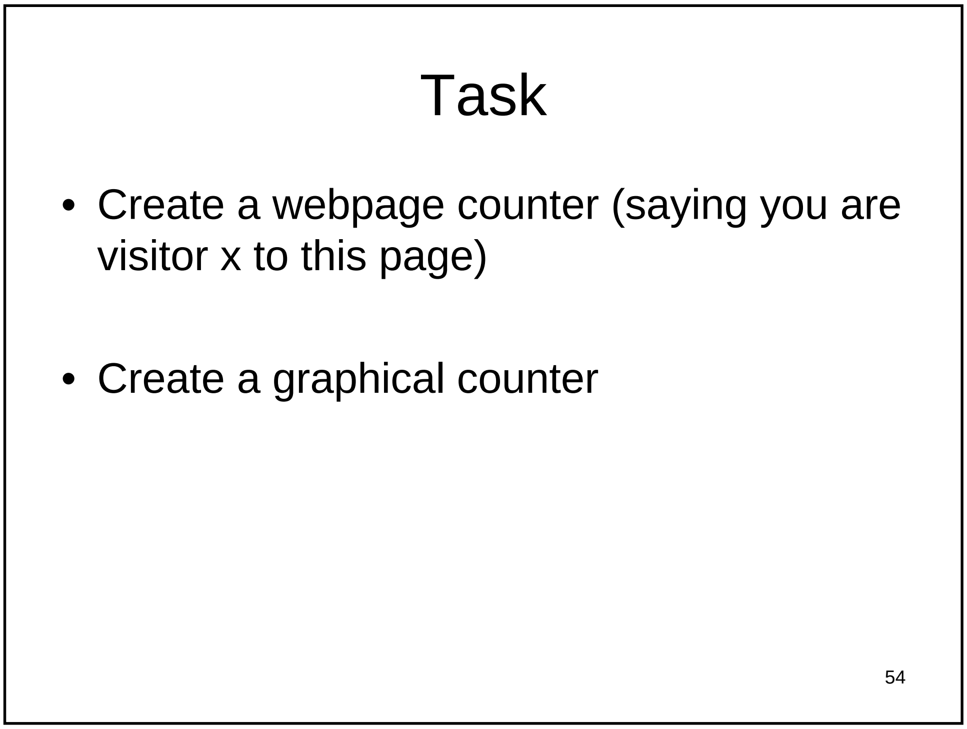
Perl debugger

- b num = breakpoint at line num
- c = run until next breakpoint
- d num = delete breakpoint at line num
- X examine all variables



Perl Debugger

- Demo of perl debugger



Task

- Create a webpage counter (saying you are visitor x to this page)

- Create a graphical counter

Wednesday is LAB!

- See you in the lab
- Need to show up in person to get lab, can stay or work offline
- Will be running lab during class session