CS3157: Advanced Programming

Lecture #4

Sept 19

Shlomo Hershkop

shlomo@cs.columbia.edu

Outline

- Feedback
- CGI
- HTML
- CGI & Perl
- Perl Debugger

- Reading:
  - Regular expressions
  - File handling

Feedback from last class

- Speed of material
- SPEAK UP!!!
  - If you don’t follow something
  - If I am going too fast
- Practical stuff

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

Perl stuff

- $|
  will turn off output buffering
- Regular expressions
  - [^something]
  - Is the negation of something in the pattern
  - /[0-9/
  - /[^0-9]

www

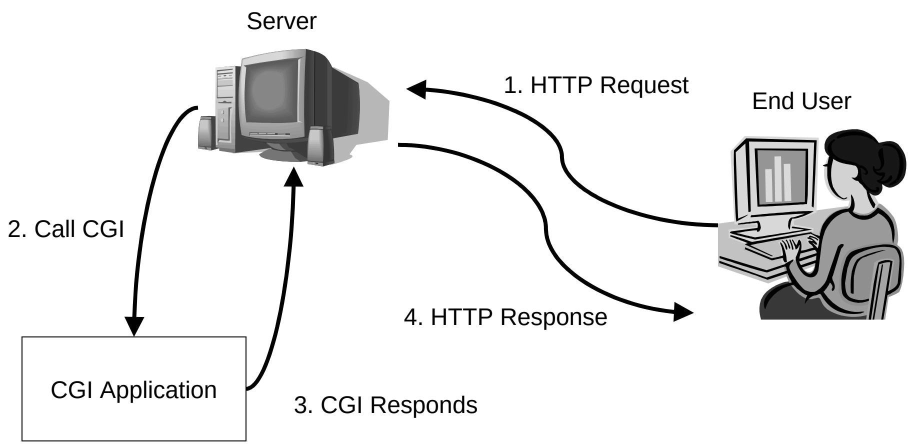
- Driven by http
- Technical overview
  - Servers serve http request
  - Clients browsers issue requests

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:

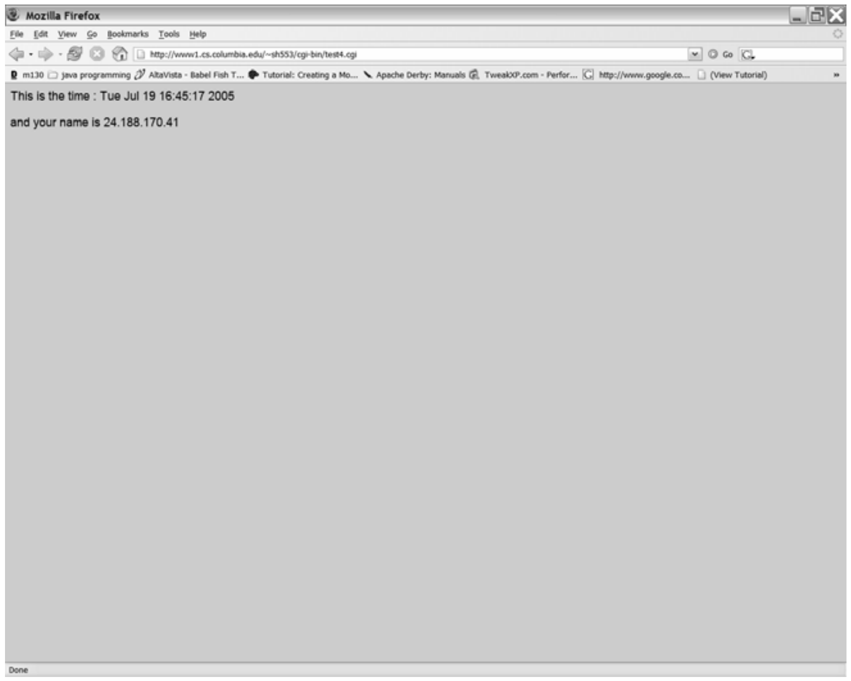


Remember

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

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

and your name is 24.188.170.41
```

Google.com

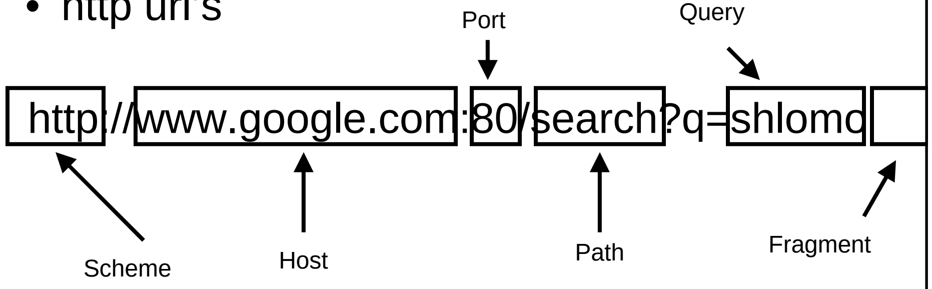
- http://www.google.com/search?q=shlomo

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

HTML

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



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

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

Some CGI Environmental Variables

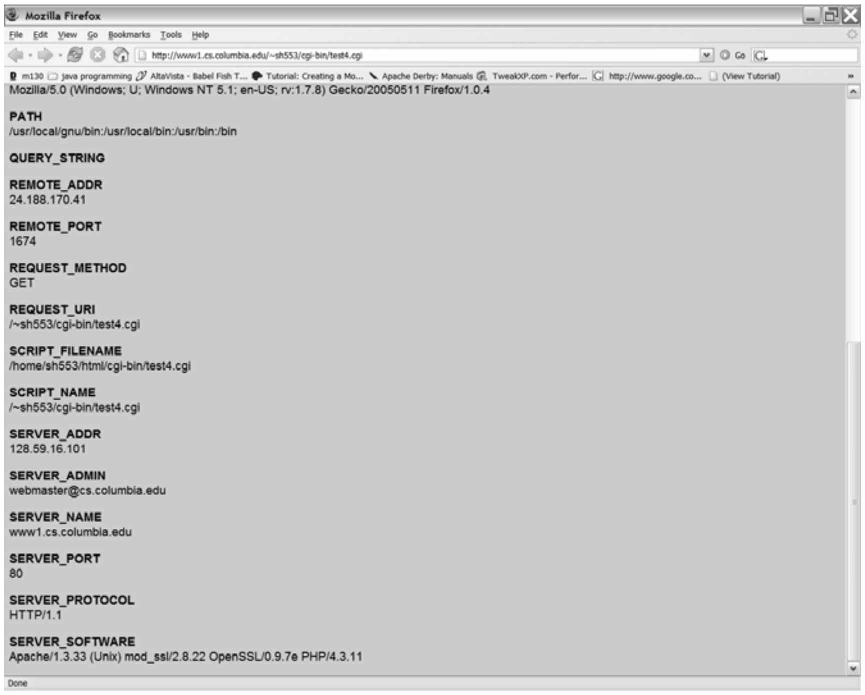
- CONTENT_LENGTH
  - Length of data passed to cgi
- CONTENT_TYPE
- QUERY_STRING
- REMOTE_ADDR
  - Ip address of client
- REQUEST_METHOD
- SCRIPT_NAME
- SERVER_PORT
- SERVER_NAME
- SERVER_SOFTWARE
- HTTP_FROM
- HTTP_USER_AGENT
- HTTP_REFERER
- HTTP_ACCEPT

Problem

- How can we print out all the environment variables ?

Example

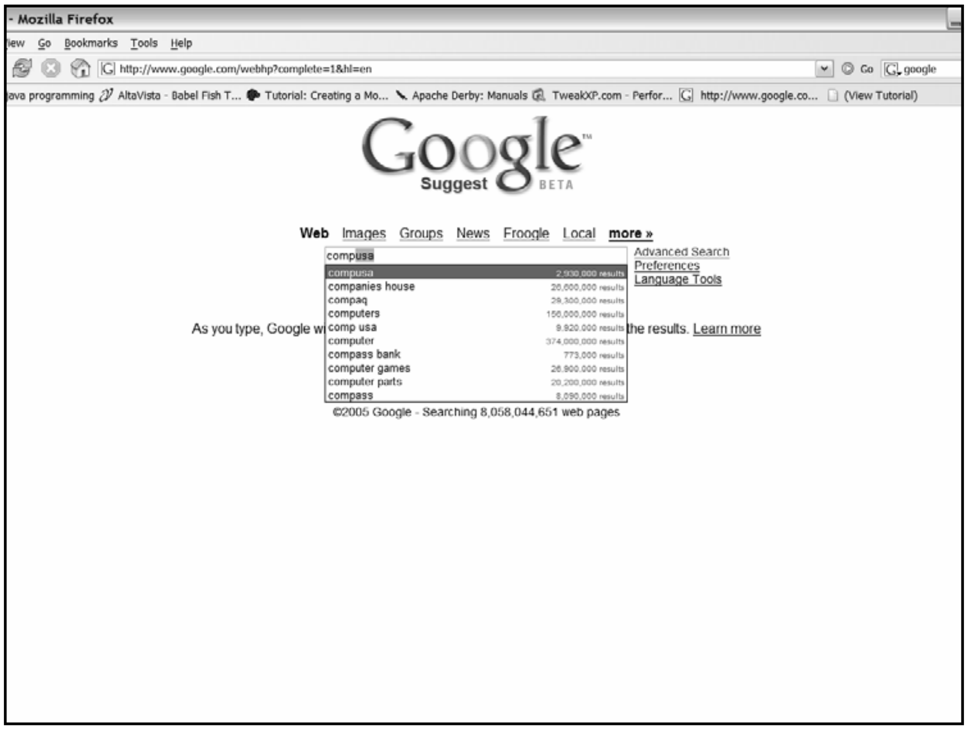
```
#!/usr/local/bin/perl

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};
}
```


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
```

Forms

- Collect data
  - Registration
  - Payment
  - Surveys
- Commands
  - Possible choice combination
  - Actions
- User needs to hit submit for anything to happen
- Google vs. Google suggest


- Mozilla Firefox
View Go Bookmarks Tools Help
http://www.google.com/webhp?complete=1&hl=en
Java programming AltaVista - Babel Fish T... Tutorial: Creating a Mo... Apache Derby: Manuals TweakXP.com - Perfor... http://www.google.co... (View Tutorial)
Google Suggest BETA
Web Images Groups News Froogle Local more »
Advanced Search
Preferences
Language Tools

compusa

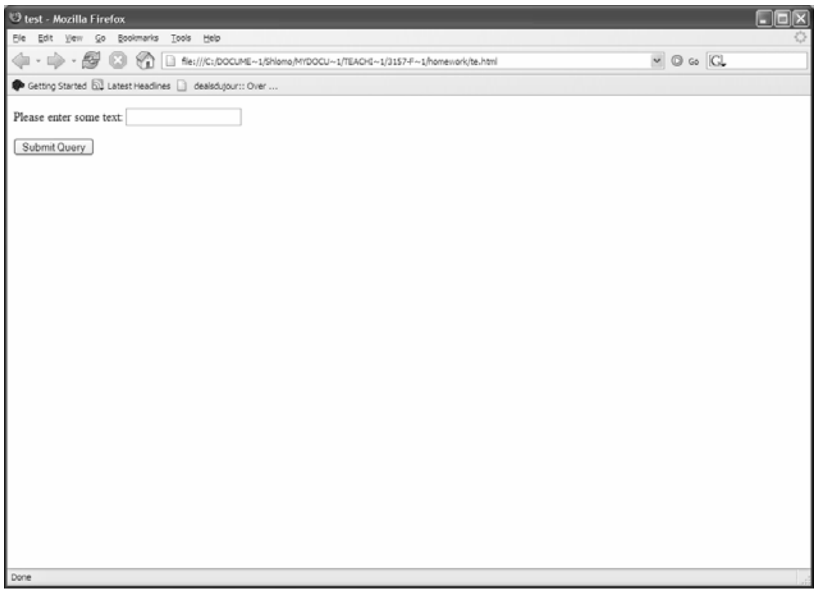
compusa	2,930,000 results
companies house	26,600,000 results
compaq	29,300,000 results
computers	156,000,000 results
comp usa	9,820,000 results
computer	374,000,000 results
compass bank	773,000 results
computer games	26,900,000 results
computer parts	20,200,000 results
compass	8,090,000 results

As you type, Google w[...] the results. Learn more

©2005 Google - Searching 8,058,044,651 web pages

Forms

```
<form action=“cgi/some.cgi” method=“GET”>
   <p> Please enter some text:
   <input type=“text” name=“string”></p>
<input type=“submit”>
</form>
```


test - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
file:///C:/DOCUME~1/Shlomo/MYDOCU~1/TEACHI~1/3157-F~1/homework/te.html
Getting Started Latest Headlines dealsdujour:: Over ...
Please enter some text:
Submit Query
Done

Interacting

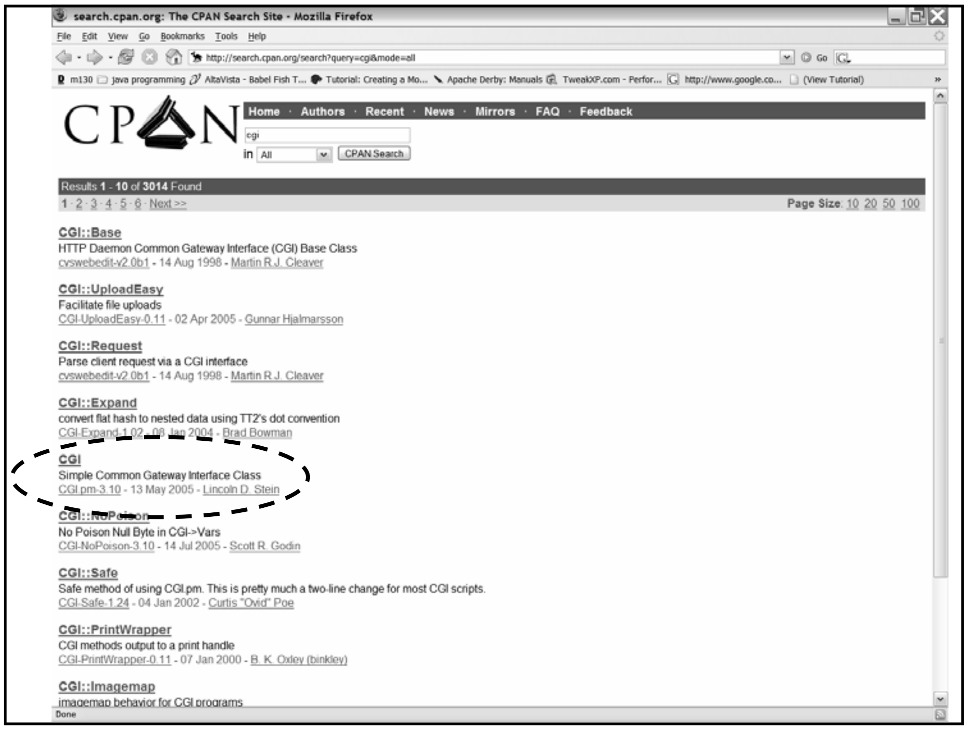
- GET
  - HTTP request directly to the cgi script
- POST
  - HTTP request in content of message
- Format:
  - Value=key separated by &
  - Space replaced by +

Decoding Form Input

1. $ENV{QUERY_STRING}
2. If($ENV{REQUEST_METHOD} eq POST)
   { read $ENV{CONTENT_LENGTH}}
3. Split pairs around &
4. Split keys and values
5. Decode URL
6. Remember key,values

Drawback

- A lot of work
- Pain if we have multiple values associated with one key
- Must be easier way…..
- CGI.pm
  - Included after 5.003_07+


search.cpan.org: The CPAN Search Site - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://search.cpan.org/search?query=cgi&mode=all
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
cgi
in All
CPAN Search
Done

Results **1** - **10** of **3014** Found

1 · 2 · 3 · 4 · 5 · 6 · Next >> **Page Size**: 10 20 50 100

CGI::Base
HTTP Daemon Common Gateway Interface (CGI) Base Class
cvswebedit-v2.0b1 - 14 Aug 1998 - Martin R.J. Cleaver

CGI::UploadEasy
Facilitate file uploads
CGI-UploadEasy-0.11 - 02 Apr 2005 - Gunnar Hjalmarsson

CGI::Request
Parse client request via a CGI interface
cvswebedit-v2.0b1 - 14 Aug 1998 - Martin R.J. Cleaver

CGI::Expand
convert flat hash to nested data using TT2's dot convention
CGI-Expand-1.02 - 08 Jan 2004 - Brad Bowman

CGI
Simple Common Gateway Interface Class
CGI.pm-3.10 - 13 May 2005 - Lincoln D. Stein

CGI::NoPoison
No Poison Null Byte in CGI->Vars
CGI-NoPoison-3.10 - 14 Jul 2005 - Scott R. Godin

CGI::Safe
Safe method of using CGI.pm. This is pretty much a two-line change for most CGI scripts.
CGI-Safe-1.24 - 04 Jan 2002 - Curtis "Ovid" Poe

CGI::PrintWrapper
CGI methods output to a print handle
CGI-PrintWrapper-0.11 - 07 Jan 2000 - B. K. Oxley (binkley)

CGI::Imagemap
imagemap behavior for CGI programs

CGI.pm

- Allows you to handle cgi in a standard format
- Can save and load key,value pairs to standard file
- Helps in creating html documents to the server by streamlining certain operations and keeping it in an object oriented design

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