

**Bo Zhang**  
**01063214**

**Task 1:**

Demonstrate that you know how to use "curl" well enough to correctly POST data to a form. Show that the HTML response that is returned is "correct". That is, the server should take the arguments you POSTed and build a response accordingly. Save the HTML response to a file and then view that file in a browser and take a screen shot.

**Curl command:** `curl -d "firstname=Roy&lastname=Zhang" http://quiet-waters-1228.herokuapp.com/echo`

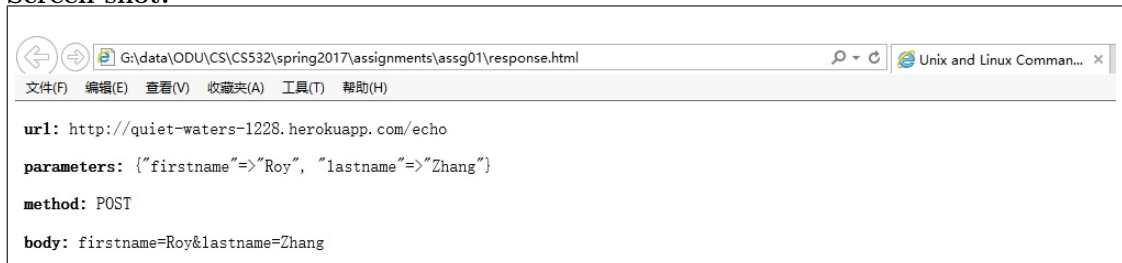
**HTML response:**

```
<!DOCTYPE html>
<html>
<head>
  <title>Unix and Linux Commands for Developers</title>
  <link href="/assets/application.css" media="all" rel="stylesheet" />
  <script src="/assets/application.js"></script>
</head>
<body>

  <p>
    <strong>url:</strong>
    <span>
      http://quiet-waters-1228.herokuapp.com/echo
    </span>
  </p>
  <p>
    <strong>parameters:</strong>
    <span>
      {"firstname"=>"Roy", "lastname"=>"Zhang"}
    </span>
  </p>
  <p>
    <strong>method:</strong>
    <span>
      POST
    </span>
  </p>
  <p>
    <strong>body:</strong>
    <span>
      firstname=Roy&lastname=Zhang
    </span>
  </p>
```

</p>  
</body>  
</html>

#### Screen shot:



**Note:** I found this website from "http://conqueringthecommandline.com/book/curl"

**Task 2:** Write a Python program that:

1. takes as a command line argument a web page
2. extracts all the links from the page
3. lists all the links that result in PDF files, and prints out the bytes for each of the links. (note: be sure to follow all the redirects until the link terminates with a "200 OK".)
4. show that the program works on 3 different URIs, one of which needs to be:  
`http://www.cs.odu.edu/~mln/teaching/cs532-s17/test/pdfs.html`

#### Algorithm:

1. Ask to input a URI
2. Open this URI and save it into an html object
3. Extracts all the links from the html object
4. Open all the links 1 by 1 and check their Content-Type from the response
5. If the Content-Type is PDF, get the URI from the response of the opened link and print it
6. Print the Content-Length from the response of the opened link

#### Results:

1. `http://www.cs.odu.edu/~mln/teaching/cs532-s17/test/pdfs.html`

```
Python 3.5.2 Shell
File Edit Shell Debug Options Window Help
Python 3.5.2 (v3.5.2:4def2a2901a5, Jun 25 2016, 22:01:18) [MSC v.1900 32 bit (Intel)] on win32
Type "copyright", "credits" or "license()" for more information.
>>>
== RESTART: G:\data\ODU\CS\CS532\spring2017\assignments\assg01\al_bzhang.py ==
Please enter a URI
http://www.cs.odu.edu/~mln/teaching/cs532-s17/test/pdfs.html
The original URI is: http://www.cs.odu.edu/~mln/pubs/ht-2015/hypertext-2015-temporal-violations.pdf
The size of the PDF is: 2184076 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/tpdl-2015/tpdl-2015-annotations.pdf
The size of the PDF is: 622981 bytes

The original URI is: http://arxiv.org/pdf/1512.06195
The final URI is: https://arxiv.org/pdf/1512.06195.pdf
The size of the PDF is: 1748961 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/tpdl-2015/tpdl-2015-off-topic.pdf
The size of the PDF is: 4308768 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/tpdl-2015/tpdl-2015-stories.pdf
The size of the PDF is: 1274604 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/tpdl-2015/tpdl-2015-profiling.pdf
The size of the PDF is: 639001 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/jcdl-2014/jcdl-2014-brunelle-damage.pdf
The size of the PDF is: 2205546 bytes

The original URI is: http://bit.ly/1ZDatNK
The final URI is: http://www.cs.odu.edu/~mln/pubs/jcdl-2015/jcdl-2015-temporal-intention.pdf
The size of the PDF is: 720476 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/jcdl-2015/jcdl-2015-mink.pdf
The size of the PDF is: 1254605 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/jcdl-2015/jcdl-2015-arabic-sites.pdf
The size of the PDF is: 709420 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/jcdl-2015/jcdl-2015-dictionary.pdf
The size of the PDF is: 2350603 bytes

>>> |
```

2. <http://www.cs.odu.edu/~mln/>

```
Python 3.5.2 Shell
File Edit Shell Debug Options Window Help
Python 3.5.2 (v3.5.2:4def2a2901a5, Jun 25 2016, 22:01:18) [MSC v.1900 32 bit (Intel)] on win32
Type "copyright", "credits" or "license()" for more information.
>>>
== RESTART: G:\data\ODU\CS\CS532\spring2017\assignments\assg01\al_bzhang.py ==
Please enter a URI
http://www.cs.odu.edu/~mln/
The original URI is: http://www.cs.odu.edu/~mln/cv.pdf
The size of the PDF is: 363963 bytes

The original URI is: http://www.cs.odu.edu/~mln/nsf-cv-2014.pdf
The size of the PDF is: 88700 bytes

The original URI is: http://www.cs.odu.edu/~mln/mln-ad.pdf
The size of the PDF is: 92868 bytes

>>>
```

3. <http://www.cs.odu.edu/~mln/teaching/>

```
Python 3.5.2 Shell
File Edit Shell Debug Options Window Help
Python 3.5.2 (v3.5.2:4def2a2901a5, Jun 25 2016, 22:01:18) [MSC v.1900 32 bit (Intel)] on win
32
Type "copyright", "credits" or "license()" for more information.
>>>
== RESTART: G:\data\ODU\CS\CS532\spring2017\assignments\assg01\al_bzhang.py ==
Please enter a URI
http://www.cs.odu.edu/~mln/teaching/
The original URI is: http://www.cs.odu.edu/~mln/pubs/phd/alnoamany-phd-dissertation.pdf
The size of the PDF is: 69827914 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/phd/brunelle-phd-dissertation.pdf
The size of the PDF is: 41389855 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/phd/salaheldeen-phd-dissertation.pdf
The size of the PDF is: 10875917 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/phd/alsun-phd-dissertation.pdf
The size of the PDF is: 23475427 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/phd/cartledge-phd-dissertation.pdf
The size of the PDF is: 20240922 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/phd/klein-phd-dissertation.pdf
The size of the PDF is: 7873533 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/phd/smith-phd-dissertation.pdf
The size of the PDF is: 4018713 bytes

The original URI is: http://www.harding.edu/fmccown/pubs/lazy-preservation-dissertation.pdf
The size of the PDF is: 4301157 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/ms/jones-ms-2015.pdf
The size of the PDF is: 21131710 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/ms/alam-ms-2013.pdf
The size of the PDF is: 1849955 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/ms/haq-ms-2008.pdf
The size of the PDF is: 622509 bytes

The original URI is: http://www.cs.odu.edu/~mln/pubs/ms/harrison-ms-2005.pdf
The size of the PDF is: 2094637 bytes

>>> |
```

**Task 3:** Consider the "bow-tie" graph in the Broder et al. paper (fig 9):  
<http://www9.org/w9cdrom/160/160.html>

Now consider the following graph:

```
A -> B
B -> C
C -> D
C -> A
C -> G
E -> F
G -> C
G -> H
I -> H
I -> K
```

L  $\rightarrow$  D  
M  $\rightarrow$  A  
M  $\rightarrow$  N  
N  $\rightarrow$  D  
O  $\rightarrow$  A  
P  $\rightarrow$  G

For the above graph, give the values for:

IN: 3 (M, O, P)  
SCC: 4 (A, B, C, G)  
OUT: 2 (D, H)  
Tendrils: 4 (N, L, I, K)  
Tubes: 1 (M  $\rightarrow$  N  $\rightarrow$  D)  
Disconnected: 2 (E, F)