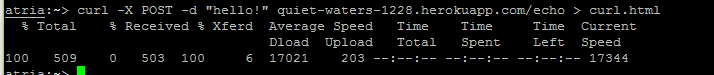
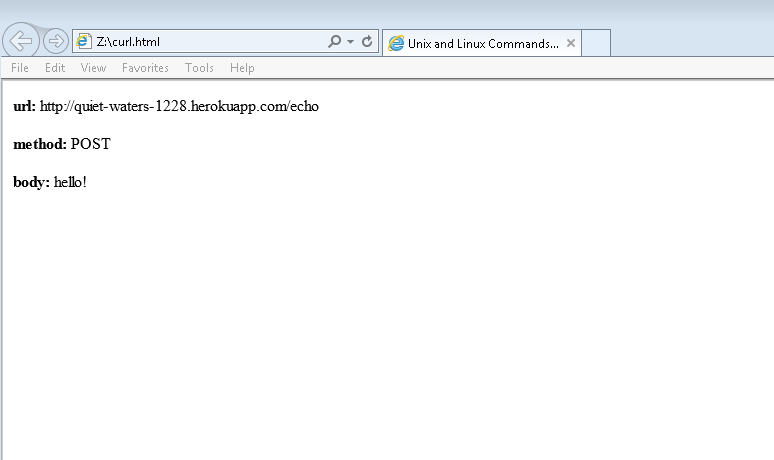
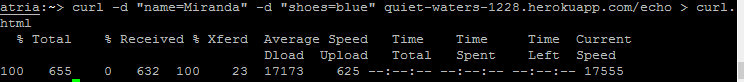
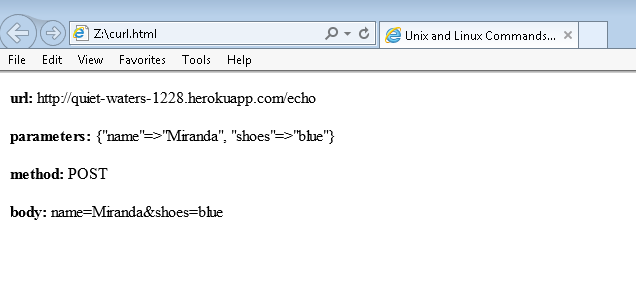
1. CURL

Found an echo server online that an online example used. Book here- <http://conqueringthecommandline.com/book/curl>





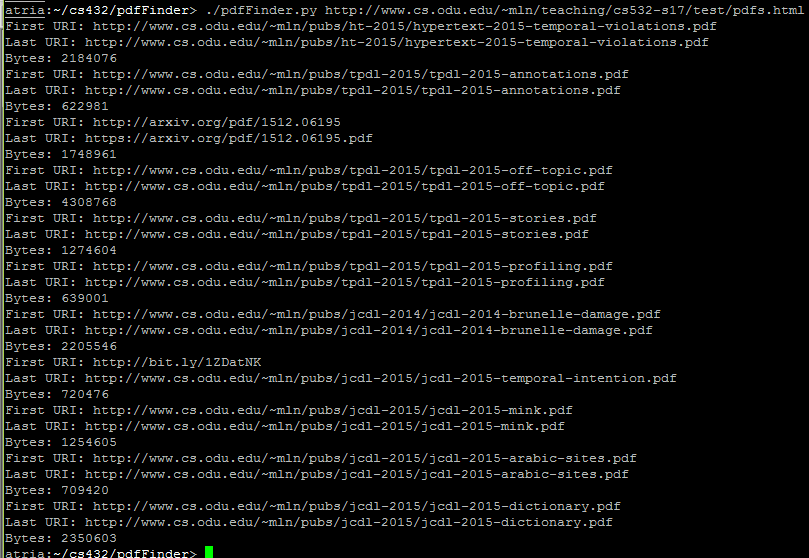




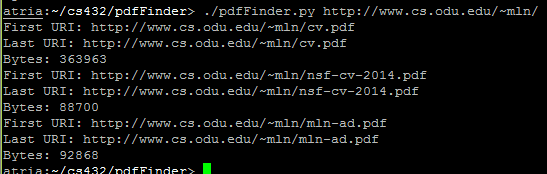
1. Python 3 program that extracts all the links from a web page and lists the pdfs with size.

Its output is as follows for 3 test cases:

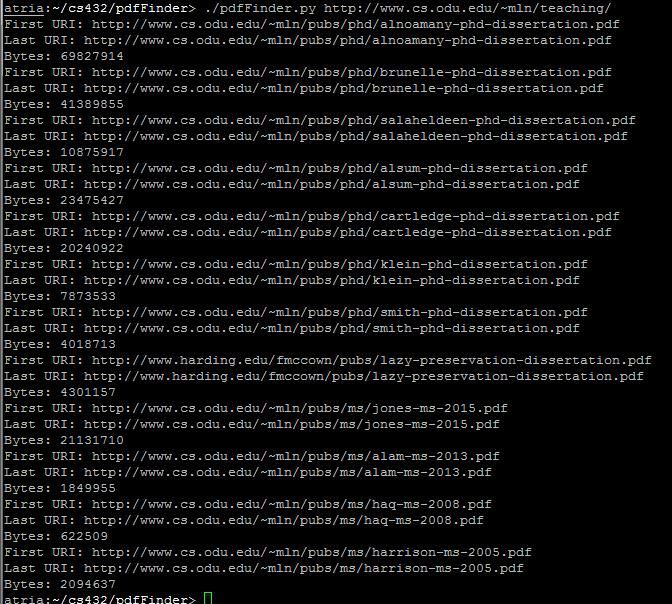
./pdfFinder.py <http://www.cs.odu.edu/~mln/teaching/cs532-s17/test/pdfs.html>



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



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



Implementation:

pdfFinder takes one command line argument, a URI. With the urllib.request library it creates a request from the inputted URI.



It then retrieves the webpage with the same library.



Using BeautifulSoup and SoupStrainer from the bs4 library, it collects all the links on the page. 

It creates a request and opens the link the same way it did the original URI. If the returned object is a pdf file according to the the Content-Type header, then the URI and size of the file is printed to the screen.



The orginal URI followed is printed out first, from the link given to the urlopen() command. The final URI is printed out from calling geturl() on the returned object to see if any redirects were followed.

3.

IN: P, O, M

SCC: A, B, C, G

OUT: D, H

Tendrils: L, I, K

Tubes: N

Disconnected: E, F