
Python For Fine Programmers

Deadline: June 18, 2009

Problem 1 (3 Points)

Write a python function/program which accepts a URL as a parameter and returns the handler to that URL as output.

Use `urllib2`.

The program should be accessible from command line.

Problem 2 (3 Points)

Write a python function which given a regular expression and a URL, returns a list of matches of the RE in the page pointed to by the URL.

Use the function from previous problem.

Problem 3 (4 Points)

Implement depth first search for directed graphs. The code for `class Graph` is provided along with the exercise.

Problem 4 (6 Points)

Write a program to combine the above exercises.

The program accepts a command line argument, a valid URL. On accepting the URL, the program creates a `Graph` with the given URL as a node.

The program finds out all the occurrences of other URLs in the given URL and adds them as new nodes to the `Graph`, and as neighbors to the node corresponding to the URL which contained them.

And recursively does the same for every node.

The program should be capable of having a control over the limit of recursion. A default limit could be 3.