

HW1 PageRank

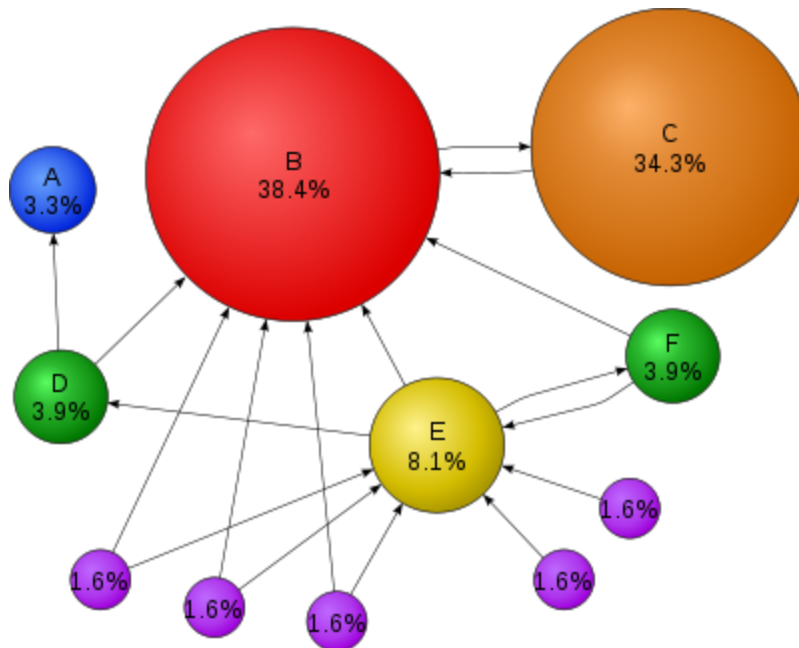
Goals

In this assignment, you will use the python programming to implement variations of the classic PageRank algorithm. You will use an adjacency matrix to represent edges and compute the PageRank scores of the nodes.

PageRank

PageRank is an algorithm developed by Sergey Brin and Larry Page that built the early foundation of the Google search engine. We will use a simplified version.

The intuition behind PageRank is that the web can be represented as a graph, with links between pages as directed edges. The importance or "popularity" of a page on the web can be determined by how many other pages link to it. But should every link be treated equally? A link from a popular site like Google's homepage should be more important than a link from an unpopular blog. We want to weigh links from more important pages more than edges links from less important edges when computing the PageRank score of a page. We need PageRank values to determine other PageRank values, and since all nodes depend on each other and the graph isn't necessarily acyclic, this becomes a chicken and egg problem!



Problems

In this assignment, you will need to solve 6 problems. Please follow the instructions in the problem files.

- Problem 1: 10 points
- Problem 2: 10 points
- Problem 3: 30 points (10 points for each subquestion)
- Problem 4: 10 points
- Problem 5: 20 points
- Problem 6: 20 points

Debugging your program using python unit tests (nosetests)

You can use the provided unit tests to debug your code. For example, to debug the code of your solution to problem 3, type:

```
nosetests test3.py
```

If the code is correct, you will pass the unit tests.

How to submit?

Please submit problem[1-6].py (six python files) in Canvas system under “HW1 assignment”.