Mini-Project 1 Writeup and Reflection

Nur Shlapobersky

In this project I created an interactive network map for a given wikipedia article and all of its connections to other articles. I used a modular approach in creating the program, separately creating python scripts that searched through wikipedia articles, created network maps, and detected keypresses. I started my searching script by initially just collecting the related links from a wikipedia page, and then added the iterative functionality so that I could take a dictionary and generate all of its contents' related articles. The other parts of my program grew in a similar fashion, starting with basic features and slowly layering on things like interactivity.

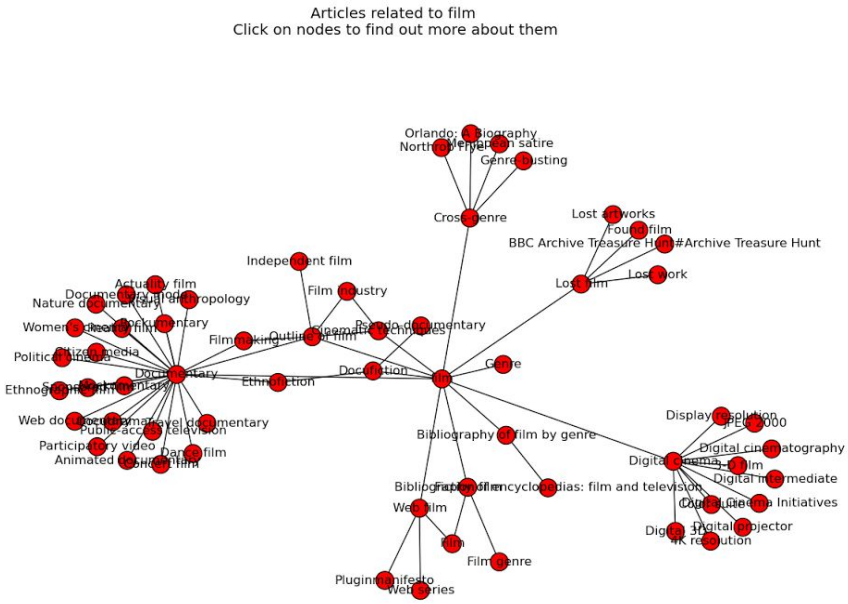
My program takes in an initial search and creates a dictionary with keys that are article titles, and values which are lists of articles related to that key. It iteratively scans through that dictionary and adds unsearched values as new keys, to a depth which the user can specify. The program then creates an interactive network map displaying all the values of the dictionary as interconnected nodes, which can then be clicked on to open corresponding web pages. I used the Pattern and Selenium libraries to interface with the internet, and the Matplotlib and Networkx libraries to create my interactive map.

By running the relationship_map.py file in the command line, the program will promt you to begin a new search with a specific depth. I used dictionaries to store article titles with a list of their related articles. A dictionary makes the most sense to use in that case, because if I had created a list of lists, I would have had to have many nested lists, iterating over which would have become far more complex than simply searching for a key in a dictionary.

The initial prompt and search output of my program:

```
nshlapo@Schizoid:~/.../MP1$ python relationship_map.py
      What subject would you like to generate a map for?Film
      What depth should the map be (start with 1 or 2)?1
        Searching through Film
      Depth 1
        Searching through Bibliography of film by genre
        Searching through Cinematic techniques
        Searching through Cross-genre
        Searching through Digital cinema
```

With an initial query of ‘Film’ and a depth of 1, my program produces the following output:



I think that my unit testing approach worked very well for this project. I was almost constantly printing outputs of my program in order to ensure that its elements were functioning properly, and as a result I never ran into large debugging issues. The greater challenge in writing this program was learning to use all the different API’s and packages and make sure they were functioning together properly. I realized about halfway through the project that I wanted the output of the program to be this interactive map, but didn’t realize how difficult creating a nice UI could be. I think that since the output of the program is this interactive map, I probably should have allocated more time to creating a cleaner output. I think the project was appropriately scoped; I felt like I had time to investigate different libraries and choose the appropriate ones, but as I said before, future work time on this project would be spent improving the interactive network graph that my program outputs.