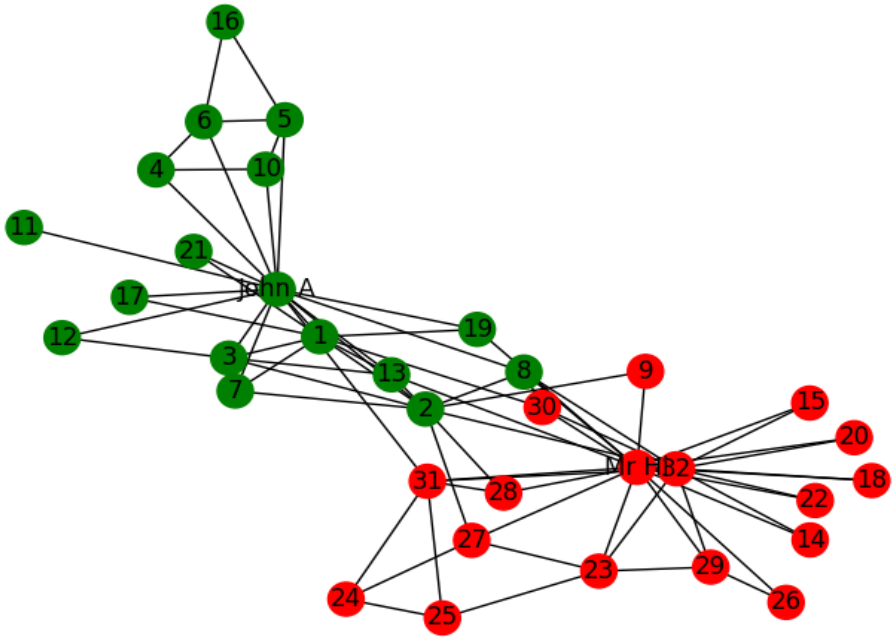
CS 432 Assignment #5
Larry Montgomery 01093132

Question 1:

The Karate Club problem is an example of the challenging nature of finding communities inside a large network. In our implementation we used the Girvan-Newman Algorithm to iterate through all the edges of the karate club social graph and find the edge in the social network that had the highest betweenness and remove it. Betweenness score for an edge represents the number of unique paths that flow through that edge connecting two other nodes in the graph. The idea is that repeatedly removing the edge involved with the most paths will be the fastest at causing discontinuity in the graph and producing the communities in the graph.

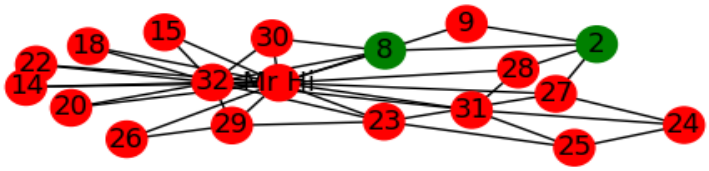
This model we have chosen to use is a good representation of social networks in reality, because it abstracts the complexity of relationships into a weighted edge between two people. This allows Us as developers to adjust the meaning of the edge leading to a flexible model. The second factor that leads me to say this represents social networks in reality, is that the model reflects how we gather information through key nodes. There are entities that connect large groups on people, and our model attempts to find those through measuring betweenness.

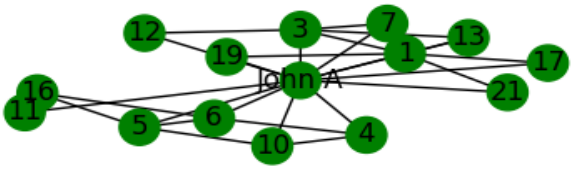
My example of the implementation started with drawing the karate graph initally. I used Matplotlib in combination with the Networkx python libraries. The picture below represents it.



Next I used the the networkx library to apply the Garvin-Newman Algorithm to find the betweeness scores for each edge. I then sorted the dictionary to retrieve the largest value from the dictionary andremove it from the edges list. Then I repeated the process until two communities we visible. I have included the edges removed and the final image for the graph after the split.

```
(('John A', '31'), 0.1272599949070537)
(('John A', '2'), 0.11924273983097515)
(('John A', '8'), 0.13782067605597018)
(('13', 'Mr Hi'), 0.14617273782105492)
(('19', 'Mr Hi'), 0.21966651886437982)
(('2', '32'), 0.1786195286195287)
(('1', '30'), 0.25601957954899124)
(('1', '2'), 0.1947415329768271)
(('2', '3'), 0.19191919191919182)
(('2', '7'), 0.25445632798573975)
(('2', '13'), 0.5080213903743315)
```





Question 2 (Extra Credit)

For the extra credit using D3.js library to plot the karate club graph before and after the split. I was able to generate the json files for each using the networkx library and was able to plot them. I ran out of time to put labels and change the colors. The plots are below.

The link: http://localhost:8000/D3_karate_graph.html
To be able to access the page you must run a simple python server from the directory the files are located. The command is " python -m SimpleHTTPServer ". This was due to accessing local files and being CORS compliant.