Henry Chang Lab 10 (I worked on this alone with assistance from algs4)

1. **How many movies has Kevin Bacon acted in?**

   Ans: 61 movies.

   Steps:
   1) Run SymbolGraph.java on movies2017.txt
   2) Type in the command line input "Bacon, Kevin (I)" to find the list of movies he starred in.
   3) Count the number of movies in the list.

2. **How many actors have co-starred with Kevin Bacon?**

   Ans: 3394 actors.

   Steps:
   1) Use the movie list from 1., loop through each movie that Kevin Bacon starred in and put their adjacent nodes of actors in a set.
   2) Count the size of the set and minus one so Kevin Bacon himself is not counted in.

3. **Co-Stars. (a) Find all the actors who have co-starred with Kevin Bacon at least twice. (b) What pair of actors has co-starred most often?**

   Since I found that there are in total 727375 actors connected to Bacon with a degree < 100, it would take to long (at least a few days if we assume the operation for each pair takes a microsecond) to find the best pair through looping through all the actor pairs, so instead, I use the actors that co-starred with Bacon more than twice for (b) as a demonstration that if given more time, we can still use the same technique to find the pair of actors who has co-starred most often among all actors.

   Ans: (a) 177 actors (please see coactors.txt for all their names);(b) David Struffolino and London Hall co-starred the most 22 time among these 178 actors (including Kevin Bacon).

   Steps:
   1) Same steps as 2., except that while looping through all the actors that co-starred with Bacon, check if they are already in the set of co-starred actors. If yes, save their name into another SET co-starred2. If not, add them to co-starred.
   2) The actors in co-starred2 is all the actors that have co-starred with Bacon at least twice.
   3) For each pair of actor1 and actor2 in co-starred2, find the sets of movies S1 and S2 they starred in, make a new set S the union of S1 and S2, count the size of S (the number of movies they co-starred), keep track of the maximum S size and the pair of actors that creates it.

5) After the loop ends, the pair of actors that is kept tracked of with the largest S size is the pair of actors that has co-starred most often(times).

4. **Find the Bacon # of all the actors; Who has the largest and second largest Bacon #'s?**

Ans: 6 Actors (Aliyeva, Rana; Basirqizi, Afaq; Baxsiyeva, G ⊣Isad; Rahimbayli, Firangiz; Ahmadpoor, Babak; Akhavan, Farhang) have the largest Bacon number of 6 (if the actors that are not connected to the graph Bacon # doesn't count.)

Steps:

1) Run BaconHistogram, observe from it that the max Bacon number is 6.
2) Modify the for loop in BaconHistogram (line50-56) so that when an actor's Bacon Number is 6 print his/her name out. In other words, modify the code to **if (bacon/2 >= 6 && bacon < MAX_BACON).**
3) The printed names is the list of names that have the largest and second largest Bacon #.

5. **Compute Kevin Bacon's Hollywood number and find an actor and actress with better Hollywood numbers.**

Ans: Kevin Bacon's Hollywood number is 2.61; Christopher Walken has a Bacon # 2.52 and Sharon Stone 2.58.

Steps:
1) I only consider the nodes that has a even degree of connection with Bacon, because otherwise the node is a movie and not an actor. This selection could be achieved with a additional statement that uses mod(2) in BaconHistogram. For simplicity, I also don't consider the nodes that have a degree larger than 100 with Bacon.
2) Run BaconHistogram with movies2017.txt and "Bacon, Kevin (I)" as input
3) Use the result in the histogram, then Bacon's Hollywood number is the (weighted sum of number of actors in each bucket weighted by each bucket #) divided by the total number of actors in the histogram. Note that the nodes with degree to Bacon > 100 are ignored.
4) Repeat 1) and 2) for each actor that co-starred with Kevin Bacon and keep track of the actor and actress that gives the Hollywood numbers smaller than Bacon.
5) Any actress or actor kept track of would be whom the question was looking for.

6. Discover something else interesting about Hollywood using this data set.

Fun fact: There are 727375 actors and actresses on IMDB that is connected with Bacon in the Graph.

Steps:
1) BFS the graph with Bacon as the source.
2) Loop through all the nodes in the graph, if the node has a even degree of connection with Bacon that is also <100, we count it as an actor connected to Bacon. If the node has an odd

degree of connection with Bacon that means it's a movie node since the graph is constructed this way by alg4.