4Ai.) Kevin Bacon

Name: Path-finding

Input: *graph:* an instance of class graph with edges representing the connections between people, *start_person:* a node in the graph representing the starting node, end_person: a node in the graph representing the ending node, *parents*: a map representing the mapping of nodes to parents

Output: a sequence of tuples of the path representing the current node and the attribute of that node

```
1. path← empty sequence
2. node← end_person
3. If start_person is not the same as end_person do
        A. While node is in the graph, do
                i. Append node to the end of path
                ii. node←$parents_{node}$
        B.Reverse the order of path
        C.final ← empty sequence
        D. If the length of path is 1, do
                i. Return an empty sequence
        E.For each integer, index in 0,1,..., length of path -1 do
                i. act1←$path_{\text{index}}$
                ii. act2←$path_{\text{index +1}}$
                iii. shared← attribute of the edge between act1 and act2
                iiii. Append the tuple (act1, shared) to the end of final
        F. Append the tuple($path_{\text{length of } path\text{-1}}$, empty set) to the end of final
4. Otherwise do
        A. Return the tuple (start_person, empty set)
5. Return final
```

1. The diameter of a graph is defined as the maximum distance between any two nodes in the graph. Describe a strategy using what you have done in this assignment for computing the diameter of a graph.

I would create a function that would find all the possible paths by looping from the start node to the end, creating a map of each path that's possible along the cumulative distance of the path. After, I would just run a loop in order to compare the distances, and I would return the largest length for the diameter.

2. In the Kevin Bacon Game, highly connected actors are "valuable". The most connected nodes — in this case actors — in a graph are called "hubs". Describe a strategy for computing the hubs in a graph.

I would create a histogram of the number of edges for every node. Then I would determine the hubs by using the histogram and seeing the number of edges on each node. Returning the determined hubs.

3. Just as certain actors can be identified as hubs in a movie graph, and therefore be seen as important to the structure to the graph, certain movies can also be considered more important to the structure of the graph. What criteria would you use to classify movies as such?

It could be determined in a similar manner using a histogram again, but this time for the number of times a movie is an edge. This is important because I think that, like in real life with highways vs small roads, a highway is very useful for being able to connect to a lot of nodes, meaning one movie will have a lot more possibilities for travel as it is more prevalent. Unlike a small road or a movie that only has a few connections. I would classify certain movies that are important to a graph as movies that connect at least 3 actors.

4. The `run` function uses the `simpleplot` module to plot the output of `distance_histogram` on two actors, Kevin Bacon and Stephanie Fratus, in the 5000-node subgraph. Describe the differences between these two plots; what does this tell you about these two actors?

The differences in these two graphs show how Kevin Bacon is more of a hub than Stephanie Fratus because Kevin Bacon has at most an average of 4 in distance and a minimum of 1, showing that for Kevin to reach any certain actor, he has a lot of connections that will lead him quickly to that actor. For Stephanie, however, her minimum average is 2 and her max average is 6, showing that she has fewer connections, so for any actor, it will usually take longer to reach than Kevin would.