In-class work: Degrees of separation*

One of the classic application of graphs is to find the degree of separation between two individuals in a social network. We'll discuss in terms of a popular game called the Kevin Bacon game.

Kevin Bacon is a well-known, prolific actor who appeared in a lot of movies. We assign every actor a Kevin Bacon number (KBN) as follows: Bacon himself is 0; any actor (except Bacon himself) who has been in the same movie as Kevin Bacon has a KBN of 1; every actor who does not have a KBN of 0 or 1, and has been in a movie with an actor who has a KBN of 1, gets a KBN of 2, and so forth.

For example. Meryl Street has a KBN of 1, because she appeared in *The river Wild* with Kevin Bacon. Nicole Kidman's number is 2 because she did not appear in any movie with Kevin Bacon, she was in *Days of Thunder* with Tom Cruise, and Cruise appeared in *A few good men* with Kevin Bacon.

Given an actor/actress name, the simplest version of the game is to find a sequence of movies alternating with actors connecting that actor to Kevin Bacon. For example: a movie buff might know that Tom Hanks was in *Joe versus the volcano* with Lloyd Bridges, who was in *High noon* with Grace Kelly, who was in *Dial M for murder* with Patrick Allen, who was in *The eagle has landed* with Donald Sutherland, who was in *Animal house* with Kevin Bacon. Based on this, Tom Hanks is at distance 5 from Kevin Bacon. But this is *not* Hanks' KBN: Hanks has KBN of 1, because he was in *Apollo 13* with Kevin Bacon.

Model this problem as a graph problem and describe algorithmically how you would solve the Kevin Bacon Game.

(a) Describe what the graph $G = (V, E)$ is in this case. What are the vertices and edges?

(b) Assume we get as input a file `movies.txt` from the Internet Movie Database. This file consists of lines, each line contains a movie title, followed by all actors who played in that movie. Describe how you go about building the graph corresponding to this file.

(c) What does the Kevin Bacon game correspond to on this graph? Describe how you would solve the game.

(d) Let's say you wanted to find an actor who is NOT connected to KB. How?

(e) Let's say you wanted to find all actors connected to KB. How?

(f) Let's say you were asked to determine an actor that is the "center" of Hollywood. How would you go about modeling it ?

*Based on Sedgewick and Wayne, Algorithms, 4th edition