Homework 0B

Deadline: 2018.03.25 (Sunday) 23:59

Problem: How does <u>Emma Watson</u> connected with <u>Leonardo DiCaprio</u>?

Do you believe the world is much smaller than we think? As the theory of <u>Six Degree of Separation</u> states: *Everything else in the world is six or fewer steps away from each other so that a chain of "a friend of a friend" statements can be made to connect any two people in a maximum of six steps*. If you don't, this assignment will make you can't afford not to believe this fact.

In this assignment, you are asked to discover the connection between two input actors or actresses in the word. Specifically, we provide you two files to do such so discovery. One is "movie_nodes.txt", and the other is "movie_edgesw.txt." The former contains the name of actor, along with his/her ID. The format of each line is: ID**\t**Name. For example, 9924 Emma Watson means the Emma Watson's ID is 9924. The latter contains the friendship between two actors' IDs. The format of each line is: ID**\t**ID**\t**Weight, in which Weight indicates the distance between two actors. Lower weight values refer to two actors are more familiar with each other. Here we provide you an example between ID 9924 and ID 2419, 9924 2419 0.34 means Emma Watson (9924) is acquainted with Daniel Radcliffe (2419), and their distance is 0.34.

Equipped with these two files, you will have a friendship network. Then your task is to <mark>discover the shortest</mark> chain of friendships that connects a source actor and a target actor (with the lowest sum of edge weights) input by the user. Here we provide you an example to exhibit the idea. Given the source and target actors are Angelina Jolie and Megan Fox, respectively, the answering shortest chain of friendships is

Angelina Jolie --0.01-- Jon Voight --0.11-- Transformers (film) --0.03-- Megan Fox

This chain implies that if Angelina Jolie would like to reach Megan Fox, **the best and shortest approach** is to first ask his father Jon Voight's help. Then since Jon Voight well involved in the movie Transformers, in which Megan Fox is the main actress, Jon Voight can best deliver Angelina Jolie's message to Megan Fox. In other words, the shortest distance between Emma Watson and Leonardo DiCaprio is 0.01+0.11+0.03=0.15

Sample Input & Output:

Case	I/O	The **Shortest** Chain of Friendships	Distance
1	Source	Emma Watson	**0.33**
	Target	Leonardo DiCaprio	
	Output	**Emma Watson -- Harry Potter and the Order of the Phoenix (film) -- Ralph Fiennes -- Naomi Watts -- W. Somerset Maugham -- Sean Penn -- Leonardo DiCaprio**	
2	Source	Tom Cruise	**0.48**
	Target	Jennifer Lawrence	
	Output	**Tom Cruise -- Robin Wright Penn -- Sean Penn -- Bill Murray -- George Carlin -- Ronald Reagan -- Dana Carvey -- Keanu Reeves -- Lori Petty -- Jennifer Lawrence**	
3	Source	Emma Watson	**0.27**
	Target	Johnny Depp	
	Output	**Emma Watson -- Harry Potter and the Order of the Phoenix (film) -- Daniel Radcliffe -- Ryan Phillippe -- Chris Cooper (actor) -- Tim Robbins -- Jack Black -- "Robert Downey, Jr." -- Johnny Depp**	
4	Source	Arnold Schwarzenegger	**0.25**
	Target	Meg Ryan	
	Output	**Arnold Schwarzenegger -- Charlton Heston -- James Garner -- Bruce Willis -- Annette Bening -- Meg Ryan**	
5	Source	Angelina Jolie	**0.15**
	Target	Megan Fox	
	Output	**Angelina Jolie -- Jon Voight -- Transformers (film) -- Megan Fox**	

Important Notes

The aim of this assignment is to let you practice the implementation of shortest path algorithm (Dijkstra) and be familiar with either Python's package NetworkX or R's package iGraph. Therefore, cannot use the following functions. But you can use them to verify your implementation.

NetworkX	https://networkx.github.io/documentation/networkx-1.10/reference/algorithms.shortest_paths.html
iGraph	http://igraph.org/r/doc/distances.html

How to Submit Your Homework?

You will need to submit two files. One is your Python/R code, and the other is the output file. Please name the Python file as "姓名_hw0b.py" or "姓名_hw0b.R". For example, if your name is 陳宜均, then your file name is: "陳宜均_hw0b.py" or "陳宜均_hw0b.R". In addition, please also name the output file as "陳宜均_hw0b.txt". Please upload your file to Moodle.