Consider the database schema for IMDB data you crawled for the assignment 2.

1> Call two movies as “similar” if there exist at least one actor who has been cast in both of them. Find the size of the largest set of movies where all the movies are mutually “similar”.

2> Consider the following scenario which explains your movie-watching behavior.

Initially you are familiar with only a few actors (in the extreme case, just one actor). The way you augment your knowledge of the acting community is by watching all the movies of the actor you already know and getting familiar with the entire cast of the movie. You repeat this process as long as new actors can be discovered.

We define the separation between two actors A and B as the minimum number of movies that you have to watch, starting with the movies of actor A, in order to discover actor B.

a) For all the pairs of actor in the database, find their separation. You have to create a new table with three columns, namely “actor_1”, “actor_2” and “separation” and populate it with the values you calculate.(You have to handle the cases where separation is not defined.) The table should be sorted lexicographically by “actor_1” column. The ties should be resolved by the lexicographic ordering of “actor_2” column.
b) From the table created in the aforementioned part, output “Yes” for all the pairs where actor_2 can be discovered with the initial knowledge of actor_1 only. Otherwise, output “No”. (Maintain the ordering described in the previous part)