Practical 7

Name: Xiaosheng Liang   ID: 15211913

Question 1:

The points of improvement in **wikihac.rb** :

● Searching the "table" after parsing a wikipage, since the names of movies are contained in the "table" and could be scraped by using tag "table".

● Scraping the tag "title", so that the name of movie could be got directly, and don't need to strip the name out.

● If there is nothing in the table we could scrape from the wikipage, it could be decided an unavailable page.

● Sometimes, the information of movie could not be scraped from the normal actor page. We should turn to its filmography page.

The result of scraping:

For uri3, 23 films could be got which used to scrape 15 films.

For uri5, 65 films could be got which used to scrape 21 films.

For uri6, 57 films could be got which used to scrape 31 films.

Question 2:

The way was used in the program is breadth first search, so the different way is depth first search. It considerably saves in memory, do better than breadth search.

L0: find the first film that KB is inconsiderably.

• Is RD a co-star in this film?

If yes, stop and say degree is 0

If no, proceed to next level

L1: find the first film of KB's co-star is in   (CO-1)

• Is RD a co-star in this film?

If yes, stop and say degree is 1

If no, proceed to next level

……

L5: find the first film of co-stars of CO-4

• Is RD a co-star in this film?

If yes, stop and say degree is 5

If no, turn to the second film that KB in (in L0) and do the search again.