

# 알고리즘 Term Project

2014860042 최훈준

## 1. 개요

### 1. 목적

같은 영화 제작에 참여한 배우들 끼리의 관계를 '친구' 라고 한다. 어떠한 영화 배우가 있을 때, '친구'인 영화배우의 '친구'로 연결되는 것을 '경로'라고 한다. 이 때, 어떠한 두 영화 배우 사이의 '친구'가 최소가 되는 최소 '경로'를 구한다.

### 2. 제약사항

영화 중 tv드라마는 제외한다.

동명이인의 경우 네이버 영화 코드를 참고한다.

## 2. 전략

### 1. 고려사항

노드의 수가 많다 (네이버 DB 기준 영화배우 약 30만건, 영화 약 10만건)

예상되는 경로의 길이가 길지 않다 (Six Degrees of Kevin Bacon)

Unweighted Graph

### 2. BFS (Breadth-First Search)

1. Enqueue the root node
2. Dequeue a node and find its adjacent nodes
3. If the target node is found in adjacent nodes, quit the search and return a result.
4. Otherwise, enqueue any adjacent nodes that have not yet been discovered.
5. If the queue is empty, the target node is not in the graph (failed)
6. Otherwise, repeat from Step 2.

### 3. 구현 (java)

#### 1. DB 크롤링

Naver 영화 사이트 크롤링

배우 : <http://movie.naver.com/movie/bi/pi/filmo.nhn?code=<code>>

영화 : <http://movie.naver.com/movie/bi/mi/basic.nhn?code=<code>>

jsoup Library 이용 HTML parsing

Google Developer Console의 SQL Engine 사용, MySQL에 저장

#### 2. 경로 탐색

기본 Queue (LinkedList) 대신 ArrayList와 HashSet을 이용, Depth별로 탐색

Tree를 만들면서 탐색, Target Node가 발견되면 root까지의 모든 parent 반환

#### 3. 전체 소스

mirror 1 : <http://130.211.255.108:8000/>

mirror 2 : [https://github.com/goznauk/UOS\\_Algorithm/tree/master/MovieStarGraph](https://github.com/goznauk/UOS_Algorithm/tree/master/MovieStarGraph)

### 4. 결과

#### 1. 김이안 - 이문수 (1)

김이안 - 스웨덴 세탁소 - 오상진 - 공모자들 - 이문수

#### 2. 권영팔 - 김희수 (2)

권영팔 - 여명 - 이민자 - 단종애사 - 장인한 - 유아독존 - 김희수

#### 3. 민진웅 - 황남 (2)

민진웅 - 패션왕 - 이일화 - 그리움엔 이유가 없다 - 윤일봉 - 격퇴 - 황남

4. 김수현 - 김삼화 (2)

김삼화 - 논개 - 이해룡 - 검은 도시 - 홍석연 - 수상한 그녀 - 김수현

5. 김소영 - 정덕순 (3) (감독 포함)

김소영 - 동면의 소녀 - 박그리나 - 바보 - 김정권 - 귀소 - 주일몽 - 님은 가시고 -정덕순

6. 다나 - 백설희 (2) (드라마 포함)

다나 - 논스톱3 - 최불암 - 사람의 아들 - 박암 - 자유부인 - 백설희