

# Movie-Chain-Runner Problem

Shashank Singh   Jimmy Zong

June 5, 2013

# Outline

- 1 Team Members
- 2 Problem Introduction
- 3 Benefits
- 4 Approach
  - Algorithms
  - Project Timeline
- 5 Evaluation
- 6 Qualifications
- 7 Summary

# Team Members

- Sung Uk Ryu
- Eugene Scanlon
- Shashank Singh
- Jimmy Zong

# The Problem

## The Problem

Find the “longest” list of overlapping titles in a list of movie titles.

# The Problem

## The Problem

Find the “longest” list of overlapping titles in a list of movie titles.

For Example: In the list

- Day of the Dead
- Live and Let Die
- Dead Poets' Society
- Die Another Day
- The Last Samurai

the “longest” chain is

“Live and Let Die Another Day of the Dead Poets' Society.”

# The Problem

- Equivalent to finding a Longest Simple Path in a directed graph

# The Problem

- Equivalent to finding a Longest Simple Path in a directed graph
- The Longest Path Problem is NP-Complete

## Previous Attempts

- Summer 2010 – 255 titles
- Fall 2010 – 271 titles (845 words)
- Summer 2011 – 311 titles (997 words)
- Fall 2011 – 323 titles (1030 words)
- Spring 2012 – 327 titles (1055 words)



# Benefits

Our group will gain experience

- programming in Python (and maybe C, MATLAB, or another language)
- working as a group toward a common goal
- handling and processing a large data set
- implementing graph algorithms
- designing and implementing approximation algorithms for an NP-hard problem

# Approach

- Algorithms
- Project Timeline (Gantt Chart)

# Algorithms

## 1 Brute Force

- Tried running on 16 GHC machines for 15 hours
- Constructed chain of 247 titles
- Progress slowed exponentially

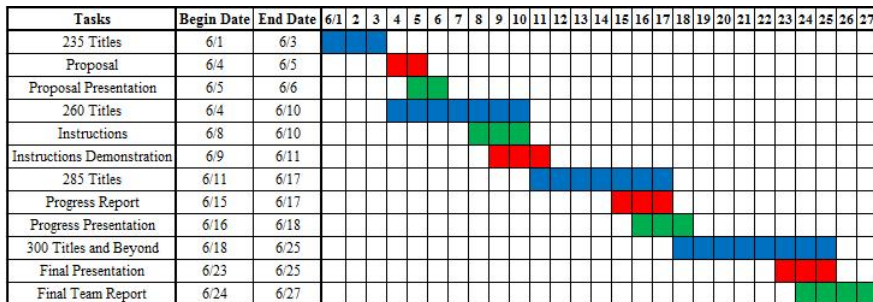
## 2 Acyclic Subgraphs

- A poly-time algorithm using topological sort is known for acyclic graphs
- Try to find acyclic subgraphs
- Too many cycles – took too long to generate subgraphs

## 3 Working backward

- Tried extending chain of 247 titles using brute force without success
- Haven't tried adding to the beginning of the chain
- Work in progress

# Gantt Chart



<b>Key:</b>	
Eugene/Sung	Red
Shashank/Yiming	Green
Everyone	Blue

# Evaluation

- 1 Length of the longest chain we find
- 2 Compare performance of a few different algorithms
- 3 Predict runtime for entire computation by solving tractable subproblems

# Qualifications

- Sung Uk Ryu
- Eugene Scanlon

# Qualifications

- Shashank Singh
  - Senior CS/Math Dual Degree
  - Completed undergrad CS/Math curricula, including courses specifically in Data Structures/Algorithms and Discrete Math (15-121, 15-211, 15-251, 15-451, 21-301, 21-484)
  - TA'd 15-211 and 15-251
  - Have extensive experience analyzing large data sets for research
- Jimmy Zong
  - Sophomore CS major
  - Completed 15-112, ...
  - Done ... projects

# Summary

- 1 Team Members
- 2 Problem Introduction
- 3 Benefits
- 4 Approach
  - Algorithms
  - Project Timeline
- 5 Evaluation
- 6 Qualifications
- 7 Summary



## Sources

- Gantt Chart created using software from the Gantt Project
  - <http://www.ganttproject.biz/> (accessed June 4, 2013)