# Final Report:
# The Movie-Chain-Runner Project

June 25, 2013

Submitted to
Thomas M. Keating
Assistant Teaching Professor
Computer Science Department
Carnegie Mellon University

Prepared and Submitted by

Shashank Singh
sss1@andrew.cmu.edu

Eugene Scanlon
escanlon@cmu.edu

# Contents

# 1 Introduction

We first introduce the problem at hand and overview our solution and its success.

## 1.1 The Movie-Chain-Runner Problem

The Movie-Chain-Runner Problem is to find the longest chain of overlapping titles in a list of movie titles, where two titles are said to overlap if some suffix of the first movie is identical to some prefix of the second movie. 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 consists of 4 titles:

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

By appropriately representing the movie list as a graph, the Movie-Chain-Runner Problem can be shown to be equivalent to the Longest Path Problem (finding the longest simple path in a directed graph). The Longest Path Problem is well known to be NP-Complete, meaning that no efficient algorithm exists to find longest paths in large graphs, including our movie title graph. A review of the literature reveals that good approximate longest paths also cannot be found efficiently in large graphs. Thus, our project is to study the movie title graph and innovate ways to find long paths in the graph.

## 1.2 Our Approach

## 1.3 Success

# 2 Approach

## 2.1 Algorithms

Ideas that worked:

We tried an algorithm that computes acyclic subgraphs of our graph and then runs a polynomial time Longest Path algorithm known for acyclic graphs. However, the number of cycles proved too large to generate the subgraphs in a reasonable amount of time, and so we abandoned this effort.

We also considered some more elaborate algorithms we found in the literature: a color-coding algorithm proposed by Alon et al.[1] and the genetic algorithms studied by Portugal

---

[1]Noga Alon , Raphael Yuster , Uri Zwick, Color-coding, Journal of the ACM (JACM), v.42 n.4, p.844-856, July 1995. (accessed June 16, 2013 at http://dl.acm.org/citation.cfm?id=210337).

| Algorithm | Longest Chain |
|---|---|
| Brute-Force | |
| Brute-Force w/ reversal | |
| Brute-Force w/ reversal and insertion | |
| DAG Poly-time | |

Table 1: Longest chain found by each algorithm

et al.[2] However, due to the complexity of implementing these algorithms and the success of our simpler algorithms, we opted not to try them.

# 3 Results

# 4 Discussion

Our group consistently met about twice a week to plan and distribute our tasks After evaluating our team's work on the project, we will reflect on lessons learned and recommendations for future groups.

## 4.1 Evaluation

We used three methods to evaluate the success of our project.

Method 1: Chain Length

Our project was a success in that the number of titles in our longest chain exceeded 300. Although we failed to establish a new record, this secondary goal was optional and not expected.

Method 2: Algorithm Comparison

Our second method of evaluation was to compare different algorithms to decide on a "best" algorithm for the problem. Although we did not test a diverse selection of algorithms, we did attempt several variants on the basic brute force algorithm and document the longest chain each variant found. Table 1 presents the results.

Method 3: Time Prediction

In order to guage the difficulty of the problem at hand, we also attempted to predict the time needed to solve the problem by exploring all possible paths in the movie graph. To do this, we first measured the time taken to solve the problem on many random subgraphs of the movie graph and used this data to fit an exponential relation between the size of the graph and the time taken for the computation. We then extrapolated this relation to conclude that performing the computation on the entire graph would take about 1.03 trillion years, approximately 100 times the lifespan of the sun. Appendix B gives a detailed discussion of the experiment and its results and validity.

---

[2]D. Portugal, C. H. Antunes, R. Rocha, "A Study of Genetic Algorithms for Approximating the Longest Path in Generic Graphs," Proc. of the IEEE SMC, pp. 2539-2544, 2010. (accessed June 16, 2013 at http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=5641920& navigation=1).

## 4.2   Lessons Learned

We learned several things about programming for large computational problems:

- Using primarily brute-force algorithms, we came quite close to (likely within 13% of) the optimal solution. Thus, although far from optimal in theory, brute-force algorithms can produce decent solutions for some NP-hard problems in practice.

- For algorithms with large asymptotic runtimes, constant factors are important - a 50% runtime reduction is very significant when the algorithm can take a week or more to run. This fact is really understated in CS classes.

- With a high level language (like Python) it is crucial to consider all available data structures and their implementations. We noticed too late that using a set rather than a list for part of our code might have sped up our program by a factor of 100.

## 4.3   Recomendations

Our team's work ethic and organization were crucial to our success. We recommend having frequent meeting (at least twice weekly) in order to coordinate work and

## 5   Sources

# ADD UTILITY SOURCES, ABSTRACT, AND COMBINE! ALSO LPP HARDNESS SOURCE!

## 6   Progress

We discuss our progress in three sections: General Progress, Status, and Projections.

## 6.1   Status

Our current longest chain (included in Appendix A) consists of 278 titles, 7 titles short of our June 17 goal of 285 titles. Since extending the movie chain has proven more difficult than we anticipated, we have decided to scale back our final goal from 300 titles to 285 titles. Since 285 titles is the necessary criterion for receiving an A on the project, the benefits of 300 titles are primarily cosmetic, and so we consider this change acceptable.

Figure 1 below shows the original Gantt chart presented in our project proposal. Figure 2 below shows our revised Gantt chart, as of June 17. The only changes are that the 300 title goal originally set for June 25 has been eliminated and the 285 title goal originally set for June 17 has been extended to June 25.
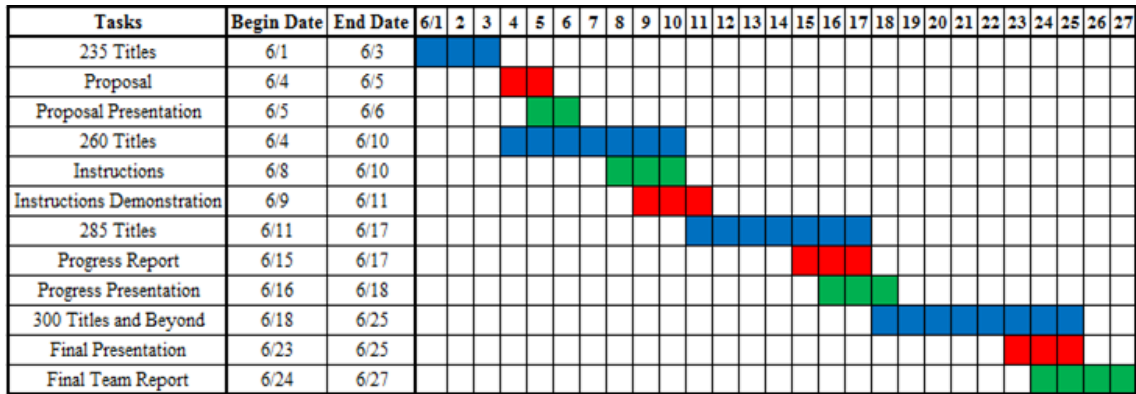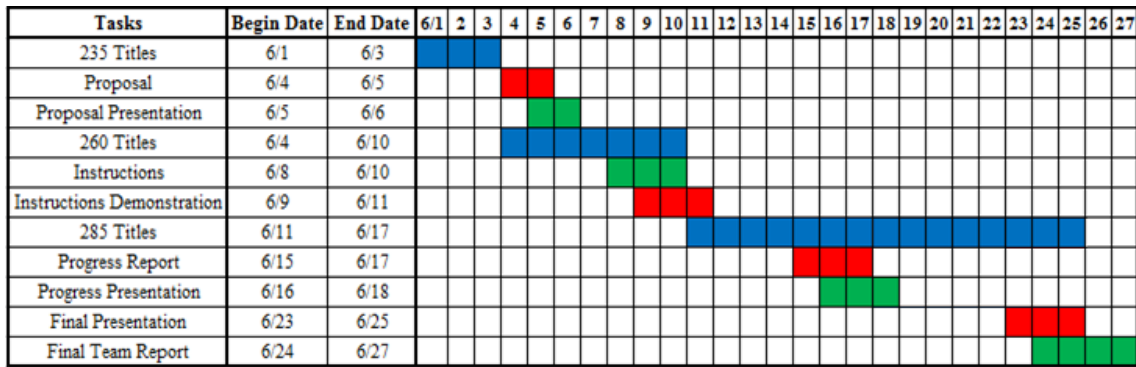
| Tasks | Begin Date | End Date | 6/1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 235 Titles | 6/1 | 6/3 | B | B | B | | | | | | | | | | | | | | | | | | | | | | | | |
| Proposal | 6/4 | 6/5 | | | | R | R | | | | | | | | | | | | | | | | | | | | | | |
| Proposal Presentation | 6/5 | 6/6 | | | | | G | G | | | | | | | | | | | | | | | | | | | | | |
| 260 Titles | 6/4 | 6/10 | | | | B | B | B | B | B | B | B | | | | | | | | | | | | | | | | | |
| Instructions | 6/8 | 6/10 | | | | | | | | G | G | G | | | | | | | | | | | | | | | | | |
| Instructions Demonstration | 6/9 | 6/11 | | | | | | | | | R | R | R | | | | | | | | | | | | | | | | |
| 285 Titles | 6/11 | 6/17 | | | | | | | | | | | B | B | B | B | B | B | B | | | | | | | | | | |
| Progress Report | 6/15 | 6/17 | | | | | | | | | | | | | | | R | R | R | | | | | | | | | | |
| Progress Presentation | 6/16 | 6/18 | | | | | | | | | | | | | | | | G | G | G | | | | | | | | | |
| 300 Titles and Beyond | 6/18 | 6/25 | | | | | | | | | | | | | | | | | | B | B | B | B | B | B | B | | | |
| Final Presentation | 6/23 | 6/25 | | | | | | | | | | | | | | | | | | | | | | | R | R | R | | |
| Final Team Report | 6/24 | 6/27 | | | | | | | | | | | | | | | | | | | | | | | | G | G | G | G |

Figure 1: Our original Gantt chart.

| Tasks | Begin Date | End Date | 6/1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 235 Titles | 6/1 | 6/3 | B | B | B | | | | | | | | | | | | | | | | | | | | | | | | |
| Proposal | 6/4 | 6/5 | | | | R | R | | | | | | | | | | | | | | | | | | | | | | |
| Proposal Presentation | 6/5 | 6/6 | | | | | G | G | | | | | | | | | | | | | | | | | | | | | |
| 260 Titles | 6/4 | 6/10 | | | | B | B | B | B | B | B | B | | | | | | | | | | | | | | | | | |
| Instructions | 6/8 | 6/10 | | | | | | | | G | G | G | | | | | | | | | | | | | | | | | |
| Instructions Demonstration | 6/9 | 6/11 | | | | | | | | | R | R | R | | | | | | | | | | | | | | | | |
| 285 Titles | 6/11 | 6/17 | | | | | | | | | | | B | B | B | B | B | B | B | B | B | B | B | B | B | | | | |
| Progress Report | 6/15 | 6/17 | | | | | | | | | | | | | | | R | R | R | | | | | | | | | | |
| Progress Presentation | 6/16 | 6/18 | | | | | | | | | | | | | | | | G | G | G | | | | | | | | | |
| Final Presentation | 6/23 | 6/25 | | | | | | | | | | | | | | | | | | | | | | | R | R | R | | |
| Final Team Report | 6/24 | 6/27 | | | | | | | | | | | | | | | | | | | | | | | | G | G | G | G |

Figure 2: Our revised Gantt chart.

| Key: | |
|---|---|
| Eugene/Sung | R |
| Shashank/Yiming | G |
| Everyone | B |

# Appendices

## A    Our Longest Movie Chain

Here, we include our current longest movie chain, with 278 titles and 885 words:

```
THE RESCUERS                        EVERYTHING RELATIVE
THE RESCUERS DOWN UNDER             RELATIVE FEAR
UNDER CAPRICORN                     FEAR STRIKES OUT
CAPRICORN ONE                       OUT OF THE PAST
ONE NIGHT STAND                     PAST MIDNIGHT
STAND IN                            MIDNIGHT RUN
IN OLD CALIFORNIA                   RUN SILENT RUN DEEP
CALIFORNIA SPLIT                    DEEP BLUE
SPLIT SECOND                        BLUE CAR
SECOND BEST                         CAR 54 WHERE ARE YOU
BEST OF THE BEST                    YOU CANT TAKE IT WITH YOU
THE BEST OF EVERYTHING              YOU LIGHT UP MY LIFE
```

MY LIFE WITHOUT ME
ME MYSELF I
I SPY
SPY HARD
HARD TIMES
TIMES SQUARE
SQUARE DANCE
DANCE WITH A STRANGER
STRANGER IN THE HOUSE
HOUSE OF DRACULA
DRACULA DEAD AND LOVING IT
IT TAKES TWO
24 7 TWENTY FOUR SEVEN
SEVEN YEARS IN TIBET
TIBET CRY OF THE SNOW LION
LION OF THE DESERT
DESERT BLUE
BLUE STEEL
STEEL DAWN
DAWN OF THE DEAD
DEAD BANG
BANG BANG YOURE DEAD
DEAD END
END OF DAYS
DAYS OF HEAVEN
HEAVEN CAN WAIT
WAIT UNTIL DARK
DARK CITY
CITY BY THE SEA
SEA OF LOVE
LOVE AND DEATH
DEATH BECOMES HER
HER MAJESTY MRS BROWN
BROWN SUGAR
SUGAR AND SPICE
SPICE WORLD
WORLD TRADE CENTER
CENTER STAGE
STAGE FRIGHT
FRIGHT NIGHT
NIGHT AND THE CITY
CITY OF JOY
JOY RIDE
RIDE THE HIGH COUNTRY
COUNTRY LIFE
LIFE IS BEAUTIFUL
BEAUTIFUL GIRLS
GIRLS GIRLS GIRLS
GIRLS JUST WANT TO HAVE FUN
FUN AND FANCY FREE
FREE WILLY
FREE WILLY 2 THE ADVENTURE HOME
HOME ALONE
ALONE IN THE DARK
THE DARK HALF

HALF LIGHT
LIGHT OF DAY
DAY FOR NIGHT
NIGHT OF THE LIVING DEAD
DEAD HEAT
HEAT AND DUST
DUST TO GLORY
GLORY ROAD
ROAD GAMES
GAMES PEOPLE PLAY NEW YORK
NEW YORK NEW YORK
NEW YORK COP
COP LAND
LAND OF THE DEAD
DEAD MAN
DEAD MAN ON CAMPUS
CAMPUS MAN
MAN OF THE HOUSE
HOUSE OF FRANKENSTEIN
FRANKENSTEIN AND THE MONSTER FROM HELL
HELL NIGHT
NIGHT FALLS ON MANHATTAN
MANHATTAN MURDER MYSTERY
MYSTERY ALASKA
ALASKA SPIRIT OF THE WILD
THE WILD ANGELS
ANGELS WITH DIRTY FACES
FACES OF DEATH
DEATH SHIP
SHIP OF FOOLS
FOOLS RUSH IN
IN COLD BLOOD
BLOOD BEACH
BEACH PARTY
PARTY GIRL
GIRL IN THE CADILLAC
CADILLAC MAN
MAN ON FIRE
FIRE IN THE SKY
SKY HIGH
HIGH CRIMES
CRIMES OF PASSION
PASSION IN THE DESERT
DESERT HEARTS
HEARTS OF DARKNESS A FILMMAKERS APOCALYPSE
APOCALYPSE NOW
NOW YOU SEE HIM NOW YOU DONT
DONT BOTHER TO KNOCK
KNOCK OFF
OFF THE BLACK
BLACK AND WHITE
WHITE LIGHTNING
LIGHTNING IN A BOTTLE
BOTTLE ROCKET
ROCKET MAN

MAN TROUBLE
TROUBLE EVERY DAY
DAY OF THE DEAD
DEAD OF NIGHT
NIGHT MOTHER
MOTHER JUGS AND SPEED
SPEED 2 CRUISE CONTROL
CONTROL ROOM
ROOM AT THE TOP
TOP GUN
GUN CRAZY
CRAZY AS HELL
HELL UP IN HARLEM
HARLEM RIVER DRIVE
DRIVE ME CRAZY
CRAZY PEOPLE
PEOPLE I KNOW
I KNOW WHAT YOU DID LAST SUMMER
SUMMER CATCH
CATCH A FIRE
FIRE ON THE MOUNTAIN
THE MOUNTAIN MEN
MEN CRY BULLETS
BULLETS OVER BROADWAY
BROADWAY DANNY ROSE
ROSE RED
RED EYE
EYE FOR AN EYE
AN EYE FOR AN EYE
EYE OF GOD
GOD IS GREAT IM NOT
NOT OF THIS EARTH
EARTH GIRLS ARE EASY
EASY MONEY
MONEY FOR NOTHING
NOTHING BUT TROUBLE
TROUBLE IN PARADISE
PARADISE ROAD
ROAD HOUSE
HOUSE PARTY
PARTY MONSTER
MONSTER IN A BOX
BOX OF MOON LIGHT
LIGHT IT UP
UP CLOSE AND PERSONAL
PERSONAL BEST
BEST MEN
MEN WITH GUNS
GUNS OF THE MAGNIFICENT SEVEN
THE MAGNIFICENT SEVEN RIDE
RIDE WITH THE DEVIL
THE DEVIL RIDES OUT
OUT COLD
COLD FEVER
FEVER PITCH

PITCH BLACK
BLACK HAWK DOWN
DOWN WITH LOVE
LOVE LIFE
LIFE OR SOMETHING LIKE IT
IT HAPPENED AT THE WORLDS FAIR
FAIR GAME
GAME OF DEATH
DEATH WISH V THE FACE OF DEATH
DEATH WISH
WISH UPON A STAR
A STAR IS BORN
BORN AMERICAN
AMERICAN HISTORY X
X THE MAN WITH THE X RAY EYES
EYES OF AN ANGEL
ANGEL BABY
BABY SECRET OF THE LOST LEGEND
LEGEND OF THE LOST
THE LOST BOYS
BOYS AND GIRLS
GIRLS WILL BE GIRLS
GIRLS OF SUMMER
SUMMER LOVERS
LOVERS AND OTHER STRANGERS
STRANGERS WHEN WE MEET
MEET JOE BLACK
BLACK LIKE ME
ME WITHOUT YOU
YOU ONLY LIVE ONCE
ONCE AROUND
AROUND THE BEND
BEND OF THE RIVER
THE RIVER WILD
WILD THINGS
THINGS TO COME
COME AND GET IT
IT HAPPENED ONE NIGHT
ONE NIGHT WITH THE KING
THE KING AND I
I WANT TO LIVE
LIVE AND LET DIE
DIE MOMMIE DIE
DIE MONSTER DIE
DIE HARD
HARD EIGHT
EIGHT AND A HALF WOMEN
WOMEN IN LOVE
IN LOVE AND WAR
WAR OF THE WORLDS
THE WORLDS FASTEST INDIAN
INDIAN SUMMER
SUMMER SCHOOL
SCHOOL OF ROCK
ROCK STAR

STAR TREK THE MOTION PICTURE

PICTURE BRIDE

BRIDE OF THE WIND

THE WIND AND THE LION

THE LION KING

KING OF THE JUNGLE

JUNGLE 2 JUNGLE

JUNGLE BOOK

BOOK OF LOVE

LOVE WALKED IN

IN GODS HANDS

HANDS ON A HARD BODY

BODY DOUBLE

DOUBLE TEAM

TEAM AMERICA WORLD POLICE

POLICE ACADEMY

POLICE ACADEMY 3 BACK IN TRAINING

TRAINING DAY

DAY OF THE WOMAN

THE WOMAN IN RED

RED RIVER

RIVER OF NO RETURN

RETURN TO HORROR HIGH

HIGH SCHOOL HIGH

HIGH SPIRITS

SPIRITS OF THE DEAD

DEAD MAN WALKING

WALKING AND TALKING

TALKING ABOUT SEX

SEX AND THE OTHER MAN

MAN OF THE YEAR

YEAR OF THE DRAGON

DRAGON SEED

SEED OF CHUCKY