# Individual Writing Exercise

Ryan Wells – 1002253w

February 26, 2013

# Introduction

*The problem is not as easy as many sports writers would have you believe... The problem is not as difficult as many mathematicians would have you believe.*

A New Property And A Faster Algorithm For Baseball Elimination
Kevin D. Wayne

## Introduction

In the competitive world of professional sport, speculation over who or what team will finish at the top of a league occurs frequently and often with arbitrary decisions based on preference and intended audience. Some sports writers have pushed this speculation to claim certain teams or individuals have absolutely no chance of finishing top of their league based on their mathematical calculations. As interesting, and sometimes controversial as this may be, naïve calculations posed by sports pundits (such as one team or individual being eliminated from success as the total of their current wins and games still to play is less than that of the current league leader) are often easily calculated or obvious, and underlying more interesting results may be found through an in-depth analysis of the league. This project aims to fulfil this analysis with a retrospective view on American Baseball Leagues from the 2011-2012 season, however any sport league with the scoring system of one point for a win, zero points for a loss and no games allowed to end with a draw will also work. This evaluation would also work with a league in progress, and arguably with more exciting implications, where scores would be added to the system as they are determined - unfortunately there is no current league in progress at the time of writing, forcing us to work retrospectively. A working knowledge of graph theory is helpful, however the specific elements utilised in this project are outlined in detail.

Complex computations such as these show their true worth in real life contexts. The information obtainable by these processes are of interest to not only professionals in the sport industry but also to fans and avid followers of teams or individuals to bookmakers and sports writers, and as such we intend to make this information readily available to anyone who wishes to view this.

We have chosen to use the American Baseball Leagues for our project due to the ease of access to accurate data and its scoring system. A scoring system of one point for a win, zero points for a loss and no draws is essential for the algorithm. We are currently only concerned with the main leagues and not playoffs or wild-card games.

## Computation

The ability to determine the elimination of a team or individual, proving that there is no possible outcome where this team or individual can finish top of a league, is a relatively easy curiosity to sate thanks to the work done by B. L.

Schwarts, proving that a maximum flow computation can determine whether a single team or individual is eliminated from a league. We are using the Ford-Fulkerson algorithm to determine our maximum flow.

A graph is constructed involving every team or individual and every match remaining in the league that does not involve the team for which we are checking the elimination status of. With the aid of a maximum flow computation on this graph we can determine the elimination status of a specific team. If a maximum flow is a saturating flow, that is to say the maximum flow that can leave the source is leaving the source, a team is not eliminated; and if a maximum flow is not saturating then a team is eliminated.

Wayne [?] proves a refinement on this idea where, if all teams or individuals in a league are ranked in a non-decreasing order by their wins and games remaining then if one team is eliminated then so are all the teams or individuals with the same or lower rank. This is a useful refinement that is utilised in our project.

## Design Decisions

Our algorithm module is given a league and a team for which we wish to determine the elimination status of, and it returns a boolean value stating this teams elimination status. This data is parsed in from a text file and the necessary objects are created with this parsed data. This information is then conveyed to the user in a graphical user interface where the user can navigate through the leagues to view scores and elimination statuses

We intend to have a desktop or web based application in which you can view a current or previous league and relevant information about this league with real time updating, where appropriate. It is hoped that through displaying this information our audience can be better informed and take an interest in the more complex methods of deduction behind our calculations.

The project intends to create three main modules: a functional and aesthetically pleasing user interface where the user can view information about teams or individuals in a certain league, a parser for a web page or RSS feed and a parser for text files so the user can get the most up to date information about their league but also view standings retrospectively, and an implementation of an algorithm that utilises the above theorems to accurately determine a team's elimination status. A user should be able to navigate through a league by week, by game and by next elimination event. The navigation will emulate a 'Time Machine' where the user can move between dates in the program and view the league status on the date they reside on. This will allow the user to view a leagues changes and inspect these events at their leisure. Each of these navigation milestones has their own benefits and it is agreed that all have merit to be in this project.

## Future Work

Further expansions to this project include determining the eliminated teams in at most log(n) maximum flow computations and providing the certificate of

elimination for all remaining teams. The certificate of elimination is a list of the team(s) that are responsible for eliminating a certain team. Another expansion to this project could be to allow a user to track their own league. The user would input the standings and teams into the program and update the scores as they came in, allowing the user to see a more localised real life example of this project. A user would also be able to create a league for their own learning where they could input results to provoke an elimination to aid in understanding how the process works. This project is not designed to be a learning tool, but a simple addition such as this can make it function as one. This would be an interesting addition to the program from the end user point of view and is currently under consideration.

The specific implementation of each module is established in the following sections of this report along with empirical evaluations and our process in developing and refinement of each module with possible improvements on each.