
Chapter One

Introduction to Ranking

It was right around the turn of the millennium when we first became involved in the field of ranking items. The now seminal 1998 paper *Anatomy of a Search Engine* used Markov chains, our favorite mathematical tool, to rank webpages. Two little known graduate students had used Markov chains to improve search engine rankings and their method was so successful that it became the foundation for their fledgling company, which was soon to become the extremely well-known search engine Google. The more we read about ranking, the more involved in the field we became. Since those early days, we've written one book *Google's PageRank and Beyond: The Science of Search Engine Rankings* and a few dozen papers on the topic of ranking. In 2009, we even hosted the first annual Ranking and Clustering Conference for the southeastern region of the U.S.

As applied mathematicians, our attraction to the science of ranking was natural, but we also believe ranking has a universal draw for the following reasons.

- The problem of ranking is elegantly simple—arrange a group of items in order of importance—yet some of its solutions can be complicated and full of paradoxes and conundrums. We introduce a few of these intriguing complications later in this chapter as well as throughout the book.
- There is a long and glorious tradition associated with the problem of ranking that dates back at least to the 13th century. The asides sprinkled throughout this book introduce some of the major and more colorful players in the field of ranking.
- It appears that today the tradition of progress on the ranking problem is reaching a peak in terms of activity and interest. One reason for the increased interest comes from today's data collection capabilities. There is no shortage of interesting and real datasets in nearly every field imaginable from the real estate and stock markets to sports and politics to education and psychology. Another reason for the increased interest in ranking relates to a growing cultural trend fostered in many industrialized nations. In America, especially, we are evaluation-obsessed, which thereby makes us ranking-obsessed given the close relationship between ranking and evaluation. On a daily basis you probably have firsthand contact with dozens of rankings. For example, on the morning drive, there is the top-10 list of songs you hear on the radio. Later you receive your quarterly numerical evaluations from your boss or teacher (and leave happy because you scored in the 95th percentile.) During your lunch break, you peek at the sports section of the newspaper and find the rankings of your favorite team. A bit later during your afternoon break, you hop online to check the ranking of your fantasy league team. Meanwhile at home in your mailbox sits the

next Netflix movie from your queue, which arrived courtesy of the postal system and the company's sophisticated ranking algorithm. Last and probably most frequent are the many times a day that you let a search engine help you find information, buy products, or download files. Any time you use one of these virtual machines, you are relying on advanced ranking techniques. Ranking is so pervasive that the topic even appeared in a recent xkcd comic.

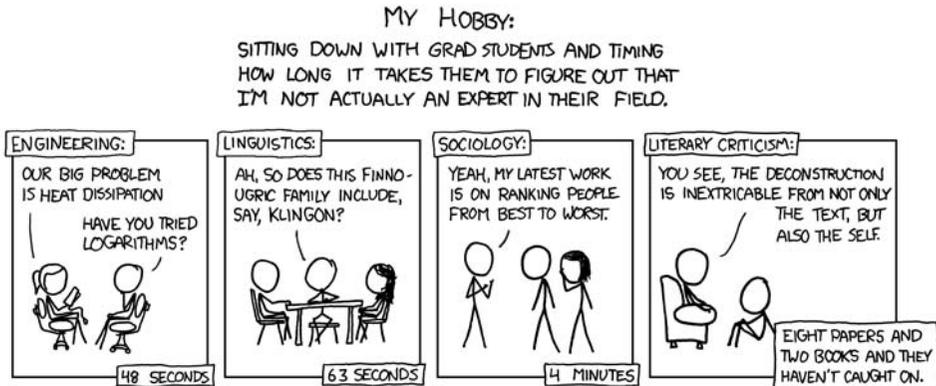


Figure 1.1 xkcd comic pertaining to ranking

Several chapters of this book are devoted to explaining about a dozen of the most popular ranking techniques underlying some common technologies of our day. However, we note that the ranking methods we describe here are largely based in matrix analysis or optimization, our specialties. Of course, there are plenty of ranking methods from other specialties such as statistics, game theory, economics, etc.

- Netflix, the online movie rental company that we referred to above, places such importance on their ability to accurately rank movies for their users that in 2007 they hosted a competition with one of the largest purses ever offered in our capitalistic age. Their Netflix Prize awarded \$1 million to the individual or group that improved their recommendation by 10%. The contest captivated us as it did many others. In fact, we applauded the spirit of their competition for it hailed of the imperial and papal commissions from a much earlier time. Some of the greatest art of Italy was commissioned after an artist outdid his competition with a winning proposal to a sponsored contest. For example, the famous baptistery doors on the Cathedral in Florence were done by Lorenzo Ghiberti after he won a contest in 1401 that was sponsored by the Wool Merchants' Guild. It is nice to see that the mathematical problem of ranking has inspired a return to the contest challenge, wherein any humble aspiring scientist has a chance to compete and possibly win. It turns out that the Netflix Prize was awarded to a team of top-tier scientists from BelKor (page 134). Much more of the Netflix story is told in the asides in this book.
- Even non-scientists have an innate connection to ranking for humans seem to be psychologically wired for the comparisons from which rankings are built. It is well-known that humans have a hard time ranking any set of items greater than size 5. Yet, on the other hand, we are particularly adept at pair-wise comparisons. In his bestseller *Blink*, Malcolm Gladwell makes the argument that snap judgments made

in the blink of an eye often make the difference between life and death. Evolution has recognized this pattern and rewarded those who make quick comparisons. In fact, such comparisons occur dozens of times a day as we compare ourselves to others (or former versions of ourselves). I look younger than her. She's taller than I am. Jim's faster than Bill. I feel thinner today. Such pair-wise comparisons are at the heart of this book because every ranking method begins with pair-wise comparison data.

- Lastly, while this final fact may be of interest to just a few of you, it is possible now to get a Ph.D. in ranking sports. In fact, one of our students, Anjela Govan, recently did just that, graduating from N. C. State University with her Ph.D. dissertation on *Ranking Theory with Application to Popular Sports*. Well, she actually passed the final oral exam on the merit of the mathematical analysis she conducted on several algorithms for ranking football teams, but it still opens the door for those extreme sports fans who love their sport as much as they love data and science.

Social Choice and Arrow's Impossibility Theorem

This book deals with ranking items and the classes of items can be as diverse as sports teams, webpages, and political candidates. This last class of items invites us to explore the field of social choice, which studies political voting systems. Throughout history, humans have been captivated by voting systems. Early democracies, such as Greece, used the standard plurality voting system, in which each voter submits just one lone vote for their top choice. The advantages and disadvantages of this simple standard have since been debated. In fact, French mathematicians Jean-Charles de Borda and Marie Jean Antoine Nicolas Caritat, the marquis de Condorcet, were both rather vocal about their objections to plurality voting, each proposing their own new methods, both of which are discussed later in this book—in particular see the aside concerning BCS rankings on page 17 and the discussion on page 165.

Plurality voting is contrasted with preference list voting in which each voter submits a *preference list* that places candidates in a ranked order. In this case, each voter creates a rank-ordered list of the candidates. These ranked lists are somehow aggregated (see Chapters 14 and 15 for more on rank aggregation) to determine the overall winner. Preferential voting is practiced in some countries, for instance, Australia.

Democratic societies have long searched for a perfect voting system. However, it wasn't until the last century that the mathematician and economist Kenneth Arrow thought to change the focus of the search. Rather than asking, "What is the perfect voting system?", Arrow instead asked, "Does a perfect voting system exist?"

Arrow's Impossibility Theorem

The title of this section foreshadows the answer. Arrow found that his question about the existence of a perfect voting system has a negative answer. In 1951, as part of his doctoral dissertation, Kenneth Arrow proved his so-called Impossibility Theorem, which describes the limitations inherent in any voting system. This fascinating theorem states that no voting system with three or more candidates can simultaneously satisfy the following four very common sense criteria [5].



K. Arrow

1. Arrow's first requirement for a voting system demands that every voter be able to rank the candidates in any arrangement he or she desires. For example, it would be unfair that an incumbent should automatically be ranked in the top five. This requirement is often referred to as the *unrestricted domain* criterion.
2. Arrow's second requirement concerns a subset of the candidates. Suppose that within this subset voters always rank candidate A ahead of candidate B. Then this rank order should be maintained when expanding back to the set of all candidates. That is, changes in the order of candidates outside the subset should not affect the ranking of A and B relative to each other. This property is called the *independence of irrelevant alternatives*.
3. Arrow's third requirement, called the *Pareto principle*, states that if all voters choose A over B, then a proper voting system should always rank A ahead of B.
4. The fourth and final requirement, which is called *non-dictatorship*, states that no single voter should have disproportionate control over an election. More precisely, no voter should have the power to dictate the rankings.

This theorem of Arrow's and the accompanying dissertation from 1951 were judged so valuable that in 1972 Ken Arrow was awarded the Nobel Prize in Economics. While Arrow's four criteria seem obvious or self-evident, his result certainly is not. He proves that it is impossible for any voting system to satisfy all four common sense criteria simultaneously. Of course, this includes all existing voting systems as well as any clever new systems that have yet to be proposed. As a result, the Impossibility Theorem forces us to have realistic expectations about our voting systems, and this includes the ranking systems presented in this book. Later we also argue that some of Arrow's requirements are less pertinent in certain ranking settings and thus, violating an Arrow criterion carries little or no implications in such settings.

Small Running Example

This book presents a handful of methods for ranking items in a collection. Typically, throughout this book, the items are sports teams. Yet every idea in the book extends to any set of items that needs to be ranked. This will become clear from the many interesting non-sports applications that appear in the asides and example sections of every chapter. In order to illustrate the ranking methods, we will use one small example repeatedly. Because

sports data is so plentiful and readily available our running example uses data from the 2005 NCAA football season. In particular, it contains an isolated group of Atlantic Coast Conference teams, all of whom played each other, which allows us to create simple traditional rankings based on win-loss records and point differentials. Table 1.1 shows this information for these five teams. Matchups with teams not in Table 1.1 are disregarded.

Table 1.1 Game score data for a small 5-team example

	Duke	Miami	UNC	UVA	VT	Record	Point Differential
Duke		7-52	21-24	7-38	0-45	0-4	-124
Miami	52-7		34-16	25-17	27-7	4-0	91
UNC	24-21	16-34		7-5	3-30	2-2	-40
UVA	38-7	17-25	5-7		14-52	1-3	-17
VT	45-0	7-27	30-3	52-14		3-1	90

This small example also allows us to emphasize, right from the start, a linguistic pet peeve of ours. We distinguish between the words *ranking* and *rating*. Though often used interchangeably, we use these words carefully and distinctly in this book. A *ranking* refers to a rank-ordered list of teams, whereas a *rating* refers to a list of numerical scores, one for each team. For example, from Table 1.1, using the elementary ranking method of win-loss record, we achieve the following ranked list.

$$\begin{array}{l} \text{Duke} \\ \text{Miami} \\ \text{UNC} \\ \text{UVA} \\ \text{VT} \end{array} \begin{pmatrix} 5 \\ 1 \\ 3 \\ 4 \\ 2 \end{pmatrix},$$

meaning that Duke is ranked 5th, Miami, 1st, etc. On the other hand, we could use the point differentials from Table 1.1 to generate the following rating list for these five teams.

$$\begin{array}{l} \text{Duke} \\ \text{Miami} \\ \text{UNC} \\ \text{UVA} \\ \text{VT} \end{array} \begin{pmatrix} -124 \\ 91 \\ -40 \\ -17 \\ 90 \end{pmatrix}.$$

Sorting a rating list produces a ranking list. Thus, the ranking list, based on point differentials, is

$$\begin{array}{l} \text{Duke} \\ \text{Miami} \\ \text{UNC} \\ \text{UVA} \\ \text{VT} \end{array} \begin{pmatrix} 5 \\ 1 \\ 4 \\ 3 \\ 2 \end{pmatrix},$$

which differs only slightly from the ranking list produced by win-loss records. Just how much these two differ is a tricky question and one we postpone until much later in the book. In fact, the difference between two ranked lists is actually a topic to which we devote an entire chapter, Chapter 16.

In closing, we note that every rating list creates a ranking list, but not vice versa. Further, a ranking list of length n is a permutation of the integers 1 to n , whereas a rating list contains real (or possibly complex) numbers.

Ranking vs. Rating

A *ranking* of items is a rank-ordered list of the items. Thus, a ranking vector is a permutation of the integers 1 through n .

A *rating* of items assigns a numerical score to each item. A rating list, when sorted, creates a ranking list.

ASIDE: An Unranked List of Famous Rankings

This book is sprinkled with examples taken from a variety of contexts and applications. In these examples, we assemble ratings, and hence rankings, and compare them to some famous standard lists. It is perhaps ironic that we begin this book on ranking with an unranked list. Nevertheless, the following list, which is in no particular order, provides a brief introduction to several standard lists that you will encounter often in this book. This unranked list also demonstrates the range of settings for which rankings are useful.

- **PageRank:** Google made this ranking famous with the success they have had using their PageRank algorithm to rank webpages [15, 49]. There is much more on this famous ranking in Chapter 6.
- **HITS:** This algorithm for ranking webpages [45, 49] was invented around the same time as the PageRank algorithm. It is currently part of the search engine Ask.com. Our Offense–Defense (or OD) method described in Chapter 7 was motivated by concepts central to HITS.
- **BCS:** College football teams are ranked by their BCS rating. BCS (Bowl Championship Series) ratings help determine which teams are invited to which bowl games. Chapters 2 and 3 cover two rating methods used by the BCS.
- **RPI:** College basketball teams are ranked by their RPI (Rating Percentage Index), which is used to determine which teams are invited to the March Madness tournament.
- **Netflix and IMDb:** Movie fans can consult the ranked lists of the best movies of all-time to determine movie rental choices. Both Netflix, the novel postal system for renting movies online, and Internet Movie Database (IMDb) provide ranked lists of the highest rated movies. The asides on pages 25 and 95 use the Netflix Prize dataset to rank movies.
- **Human Development Index:** The United Nations ranks countries by their Human Development Index (HDI), which is a comparative measure of life expectancy, literacy, education, and standard of living. The HDI is used to make decisions regarding aid to under-developed countries.
- **College Rankings:** *US News* publishes an annual ranking of colleges and universities in the United States. The aside on page 77 contains more on these rankings and their calculation.
- **Chess players:** FIDE, the international federation for chess, uses the Elo method of Chapter 5 to rank hundreds of thousands of chess players worldwide.

- Food chain: Species of animals feed on and are food for other species. Thus, a directed graph showing the dominance relationships between different species can be built. This graph can then be analyzed to answer questions such as who is top of the food chain or which species is the linchpin of the entire chain.
- Archaeology: The aside on page 192 shows how ranking methods are used to put artifacts in relative order based on their dig depths at multiple locations.
- Legal cases: The trail of citations from case to case to case leaves valuable information that enables a ranking of legal cases. In fact, current legal search engines use this technology to return a ranked list of cases in response to a paralegal's query on a particular topic.
- Social Networks: The link structure in a social network like Facebook can be used to rank members according to their popularity. In this case, the Markov or OD ranking methods of Chapters 6 and 7 are particularly good choices. However, Facebook may have had its origins in the Elo rating system described in Chapter 5—see the aside on page 64. Of course, some networks, such as the links between terrorists coordinating the 9/11 attacks, carry weightier communications. The same techniques that find the most popular Facebook members can be used to identify the key figures, or central hubs, of these networks.

By The Numbers —

50,000,000 = attendance at American college football games in 2010.

— The Line Makers