CSC2421: Online and other myopic algorithms
Spring 2021

Allan Borodin

January 21, 2021

Week 2

Todays agenda:

- I will discuss what is in Chapters 1,2,3. These chapters and the index are posted on the web page. I will be alternating between these slides as an outline, and the material in the Chapters.
- I will then again do a quick overview of the (intended) chapters. Hopefully the initial chapters will give us some sense of the scope of the topic.
- Generally, we will be presenting both positive and negative results. Often people say upper bounds (for positive) and lower bounds (for negative) results. The meaning of this is clear when dealing with minimization problems. However, when dealing with maximization problems, the conventions are not consistent so I will try to use the terms competitive ratio for an algorithm or class of algorithms (i.e., a positive result) and an inapproximation (i.e. a negative result) for an algorithm or a class of algorithms.

What's in Chapter 1

- We begin with a little history and then the chapter discusses three problems for motivation; namely the ski rental problem, the line search problem, and the paging problem.
- As I mentioned last week, much of the interest in online algorithms in TCS stems from the 1985 Sleator and Tarjan paper in which they discuss two problems, paging and list accessing. This is often cited as th beginning of competitive analysis as it argued for worst case analysis of online algorithms (i.e., competitiive analysis although it did not use that terminology).
- While the Sleator and Tarjan paper was indeed a seminal paper, one should not be surprised that there was previous work that can be seen as competitve analysis.
- In particular, last week we called attention to Graham's 1966 and 1969 papers on the makespan problem). The emphasis there was more on simple worst case approximation algorithms such as the greedy LPT algorithm. Graham asked if LPT was the best "list processing algorithm".

What's in Chapter 1 continued: more history

- Johnson in his 1974 thesis asked if First Fit or Best Fit (to be discussed in Chapter 2) was the best online algorithm for bin packing. Johnsons thesis was preceded by a 1973 conference paper by Garey,Graham and Ullman and that paper was precided by a 1971 technical report by Ullman. The culmination of all these works appears in a 1974 Johnson et al journal paper.
- In 1980, Yao explicitly considered the question as to the competitive analysis for the online bin packing problem (without using that terminology). That is, what is the best (worst case) online algorithm for bin packing.
- And even before these TCS papers, Beck (1964) and Bellman (1963) considered the analysis for the line searching problem.
- And if one extends competitive analysis to the random order model, Lindley (1961) is given credit for the analysis of the "secretary problem" although folklore traces the interest in the problem back to kings choosing wives.

What's in Chapter 1 continued: the ski rental problem

- The ski rental problem seems like a very simple stylized problem but it is an abstraction of a basic problem that led to a sequence of ideas and extensions (see Chapter 7) that make it an interesting problem.
- To be historically accurate, the ski rental problem came up as being the crux of a *phase* in a multiprocessor caching system called "snoopy caching". Snoopy caching is just one example of online problems dealing with page migration and replication in multiprocessor systems.
- The terminology "ski-rental" is not mentioned in the Karlin et al paper on snoopy caching. The terminology was later coined by Rudolf one of the authors of the snoopy caching paper.
- **Aside:** It is interesting to note the role of suggestive terminology.
- Ski rental was also one of the early problems studied with respect to randomized algorithms and in Chapter 1 we analyze ski rental from the perspective of the competitive analysis for randomized online algorithms. Chapter 3 formally introduces randomized algorithms.

Chapter 1 continued: the line search problem

- This seems like another seemingly very stylized problem. It is also sometimes refered to as the "cow-pasture problem".
- The problem has a long history dating back (independently) to Beck [1964] and Bellman [1963]. We will see that this problem is a little different than almost all the problems we consider in the text as I will explain when we get to Chapter 2.
- It has been extended to the m-ray problem where the cow is searching for a pasture somewhere in an infinite star graph with m rays emenating from the origin. The line search problem is then the m-ray problem for $m = 2$.
- The analsys of the problem has an application in a problem called the "parking problem" on the line which is a special case of min cost matching.
- The analysis introduces an idea that is used in many online algiorithms, mainly this doubling technique of "trying something for some assumed bound t and if that doesn't work doubling t and trying again".

The paging problem

Unlike the previous two problems which serve as simple abstractions for more "practical applications", the basic paging problem is studied by practitoners, albeit not in terms of the basic competitive analysis that was one of the two examples in the seminal Sleator and Tarjan paper.

- There are a number of paging algorithms, two prominent ones being LRU and FIFO. When there are k fast memory pages, both algorithms are k competitive (i.e. the competitive ratio is k) and this is optimal.
- Both are examples of *marking algorithms* and every marking algorithm is k competitive.
- In practice, LRU is known to be a much better paging algorithm than FIFO and skome vafriant of it is used in various caching systems. The fact that competitive analysis does not distinguish between these two algorithms is just one reason that competitive analysis seems inappropriate for the paging problem. But moreover, competitive analysis is a worst case measure and hence is very pessimistic especially compared to experimental real system analyis.

More on paging using competitive analysis

- It is also counter intutiive to have a result that gets worst as the number of pages increases.
- In some sense these criticisms have led to interest in going "beyond worst case analysis". This has led to alternative measures of performance as well as more realitistic models of input sequences.
- In particular, the basic competitive analysis fails to account for "locality of reference" as already studied by Denning in 1968. In a relatively recent paper, Albers et al (2005) give precise theoretical results relating to locality of reference showing that LRU is an optimal online algorithm with respect to the *page fault measure* and superior to FIFO and other marking algorithms.

What's in Chapter 2

Overview of Chapter 2

- We introduce the framnework of request-answer games for online problems. This framework will suffice for almost all the problems we consider with the major exception of navigation problems (such as the line search problemi as discussed in Chapter 1 and other navigation problems planned for Chapter 23).
- We provide the deterministic online algorithm template
- The competitive ratio of an algorithm and the competitive ratio of an online problem are defined.
  We also define a strict competitve ratio and discuss the difficulty of clarifying what we mean by an asymptotic competitve ratio.
- The ultimate goal for a given problem is to have an efficient online (or myopic) algorithm achieving a competitive ratio ρ and an asmyptotic inapproximation ρ for any online algorithm.

Chapter 2 continued

- The chapter first presents two minimization problems. One we have already considered is the makespan problem (for identical machines). Even after 50 years of consideration, we still do not know the precise competitive ratio for a given number m of machines.
- The second minimization problem is the bin packing problem which has been extensively studied and yet is also still not fully resolved. There are three natural online algorithms which are now fully understood. The early (1971,1973,1974) work by Johnson et al showed that the asymptotic ratio for the FirstFit and Best Fit algorithms is $\frac{17}{10}$. Forty years later Dosa and Sgall show that $\frac{17}{10}$ is the strict competitive ratio. Yao (1980) exhibited an online algorithm with aymptotic competitive ratio $\frac{5}{3}$ and a $\frac{3}{2}$ asymptotic inapproximation for any online algorithm.

Chapter 2 continued: maximization problems

Chapter 2 also considers the definition of the competitive ratio for maximization problems. Unfortunately, there are two ways to define the ratio, one where the ratio is at least 1 (as it is for minimization problems) and one where it is at most 1.

It would ne nice if the literature would converge on one definition but that doesn't seem to be happening. I think we are trying to use the convention that when the ratio is an absolute constant (say as we will see for the online maximum cardinaltiy bipartite matching problem in Chapter 5), we will state the ratio as a fraction. When the ratio depends on some input parameters as it does in Chapter 2, we will use ratios that are greater than great than 1.

The chapter considers two related "financial problems", namely the *time-series problem* and the *one-way trading problem*. Both problems are parameterized by L and U, the minium and maximum exchange rates.

Time-series search and one-way trading

When L, U are known, we then obviously also know $\phi = \frac{U}{L}$ (but not conversely). The deterministic competitive ratios for both problems are stated in terms of ϕ. Without loss of generality lets take $U \geq L \geq 1$.

- In the time series problem we are trying convert all of our assets (say in US $ tho Canadian $) in one day. Here we can achieve competitive ratio $\sqrt{\phi}$.
- In the one way trading problem, we are allowed to gradually trade the currency until some day when the trading ends. In this problem we can achieve competitive ratio $c(\phi)\log_2 \phi$ where $c(\phi) \to 1$ as $\phi \to \infty$.
- The algorthm uses another common idea in algorithms, namely partitioning the input space into a number of classes. That idea is used in the design and analysis of bin packing algorithms. The idea of a logarithmic number classes is also a commonly used idea.

End of week 2 outline

We ended the week 2 discussion just mentioning the bin packing, time-series search and one-way trading problems.

Next week we will briefly consider these three maximization problems as they appear in Chapter 2.

We will then discuss the material in Chapter 3 (randomized online algorithms).

Then I would like to see what possible topics are of interest for the course reading requirement.