Chapter 12: Lower Bounds for Online Algorithms

CMSC 858M: Algorithmic Lower Bounds: Fun with Hardness
Instructors: Mohammad T. Hajiaghayi and William Gasarch
Textbook: Fun with Hardness: Algorithmic Lower bounds

Juan Luque

Spring 2021

1 Introduction

In *online* problems we consider problems where input arrives over time, and we must continuously make decisions as information arrives. Due to the incomplete information it is generally impossible to make optimal decisions. For instance, consider the following example.

Example 1. *Consider a multiple queue problems were agents a_i arrive one at a time and must be irrevocably assigned to a queue q_j upon arrival. Further, helping each agent takes time t_i and our goal is to minimize the time at which the last queue finishes.*

This queueing problem is simply load balancing problem, which is known to be NP-hard. Solving this problem optimally can be done (with enough time), however finding the optimal solution in the online setting is not possible without knowing the information of agents prior to their arrivals.

This motivates the definition for measuring the performance of online algorithms where we compare the performance of an algorithm relative to that of the optimal solution.

Definition 1 (Competitive ratio)**.** *Suppose we have a sequence of inputs σ. Knowing all of σ a priori we can solve the problem optimally for* $\mathrm{OPT}(\sigma)$*. Now we will obtain the competitive ratio by optimally. Then* ALG *is said to have a competitive ratio of α if the underlying problem is a*

1. *maximization problem, and $\forall\sigma$, $\mathrm{ALG}(\sigma) \geq \alpha \cdot OPT(\sigma)$, or*
2. *minimization problem, and $\forall\sigma$, $\mathrm{ALG}(\sigma) \leq \frac{1}{\alpha} \cdot OPT(\sigma)$.*

Moreover, ALG *is said to be α-competitive.*

With this definition we have $\alpha \in (0, 1]$ such that larger α values are indicative of better worst-case online algorithms. It follows easily that $\alpha = 1$ indicates that ALG = OPT.

Note that there is some inconsistency with the definition of competitive ratio in the literature. Sometimes it is defined as

2 Yao's lemma

Yao's lemma is a powerful tool for proving results in the performance of randomized algorithms. Let $c(a, x)$ be the cost incurrecd by algorithm a on input x. Then Yao's lemma is as follows.

Lemma 1 (Yao's lemma)**.** $\max_{x\in\mathcal{X}} \mathbb{E}[c(a, x)] \geq \min_{a\in\mathcal{A}} \mathbb{E}[c(a, x)]$.

Thus the first expectation is taken over the choice of algorithm a and the second expectation is taken over the choice of input x. Intuitively, Yao's lemma states that the performance on the worst input (averaged over all algorithms), is worse than the performance of the best algorithm (averaged over all inputs).

3 Online matching

The study of online matching began with the seminal paper of [KVV90] which studied online bipartite matching with vertex arrivals and gave an optimal $(1 - \frac{1}{e})$-competitive algorithm. Subsequently, [BM08] resolves a slight error in a proof from [KVV90] and re-proves the problem in an intuitive manner.

An alternate definition of competitive ratio has been used in other online matching problems such as [BSSX16]. This definition considers peformance in expectation and is especially helpful when we consider stochastic graphs where edges or vertices are realized with some probability. Contrast this with the previous definition which considers the worst case performace of ALG thus it considers adversarial (worst case) arrivals.

Definition 2 (Competitive ratio; alternate definition)**.** *Define the competitive ratio α as*

$$\alpha = \frac{\mathbb{E}[ALG(\sigma)]}{\mathbb{E}[\mathrm{OPT}(\sigma)]}$$

where the expectation is taken over the inputs σ and any internal randomness of ALG*.*

Additional online matching settings studied include permutations of (1) vertex and edge arrival models, (2) weighted and unweighted settings, (3) known i.i.d. An interested reader may refer to [BSSX16].

4 Further results

1. [GKM+19] Under the adversarial edge arrival setting, randomization does not help beat the trivial $\frac{1}{2}$ competitive ratio.

2. [AM17] Role-matchmaking is a problem where we have players with skills levels as well as preferences over specific roles they would like to play (e.g., in soccer we have roles of goalkeeper, defender, midfield, or forward). This problem has immediate applications to many popular online videogames such as *League of Legends* and *Dota 2*. Then, assuming the 3SUM conjecture, role-matchmaking is intractable.

3. [CK18] Consider the problem of finding a popular matching in a graph G with n vertices. The problem is NP-complete for even n, but efficiently solvable for odd n. A matching M is considered popular if it does not lose against any matching M' in a head-to-head election where each vertex gets to vote.

4. [Hua10] A *super stable matching* is a matching that disallows blocking triples of degrees 1, 2, or 3. Moreover a blocking triple is an analogous extension of *blocking pairs* from the well known Gale-Shapley algorithm for the stable matching problem. A blocking tri Deciding whether a super stable matching exists in a circular stable matching problem with ties in the preferences is NP-complete. This is true even if all ties are of size at most 3 and they are at the front of the preference lists.

5 Chapter suggestions

1. p330 Theorem 18.4.3 uses a different definition of competitive ratio than what is given at the beginning of the chapter. I believe these are reciprocals of each other. A suggestion would be to stick with the convention being used in the Theorem since it seems to be most commonly used in current research.

2. p327: Under Yao's Lemma. The second enumerated item has $\max_{x \in A}$ but it should read $x \in X$. Same thing below in the lemma environment.

   (a) backwards quote right under the Lemma at the top of p328. Should have a "the worst input in ....

   (b) Suggested edit for the intuition of Yao's Lemma: cost under the worst input in $p \geq$ cost of the best deterministic algorithm w.r.t. p. Note the case used for the distribution p. Also, the brief explanation should mention we are comparing costs. Otherwise it isn't clear if the $\leq$ means less cost or better.

References

[AM17] Josh Alman and Dylan McKay. Theoretical foundations of team matchmaking. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, pages 1073–1081, 2017.

[BM08] Benjamin Birnbaum and Claire Mathieu. On-line bipartite matching made simple. *Acm Sigact News*, 39(1):80–87, 2008.

[BSSX16] Brian Brubach, Karthik Abinav Sankararaman, Aravind Srinivasan, and Pan Xu. Online stochastic matching: New algorithms and bounds. *Algorithmica 82, 2737–2783*, 2016.

[CK18] Ágnes Cseh and Telikepalli Kavitha. Popular matchings in complete graphs. *arXiv preprint arXiv:1807.01112*, 2018.

[GKM+19] Buddhima Gamlath, Michael Kapralov, Andreas Maggiori, Ola Svensson, and David Wajc. Online matching with general arrivals. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 26–37. IEEE, 2019.

[Hua10] Chien-Chung Huang. Circular stable matching and 3-way kidney transplant. *Algorithmica*, 58(1):137–150, 2010.

[KVV90] Richard M. Karp, Umesh V. Vazirani, and Vijay V. Vazirani. An optimal algorithm for on-line bipartite matching. *22nd Annual ACM Symposium on Theory of Computing*, 1990.