

COS 445 - PSet 1, Problem 3

Mark Watney

February 15, 2021

Problem 3: Random Preferences

We wish to prove that the expected sum of the students' ranks is $\mathbb{E}[\sum_i X_i] = O(n \log n)$.

We first show that the algorithm terminates after every university is proposed to. This is relatively straightforward, because there are both n students and n universities. There will not be an unmatched student, since a student must rank all universities and can potentially propose to all of them if their first $n - 1$ choices are already taken. Since each university takes exactly one student, there also cannot be an unmatched university since at worst a max of $n - 1$ students will match with other universities.

Additionally, the sum of the students' ranks is equal to the sum of the proposals made. Intuitively, each time a student is rejected their rank will increase by one. This is because the university they propose to in the next round will be one lower on their preference list than the previous university they proposed to.

It is given that each student's preferences are drawn uniformly at random. This means that each university has an equal chance to be at any point of a given list. By the principal of deferred acceptance, we can treat students as if we are unaware of their full preferences. During round i , for example, we can simply ask students what their i -th preference is.

In this context, there is a relation between this problem and the Coupon Collector Problem. We can couple the probability that a student proposes to a given university to the probability that a collector draws a distinct card. The student who first proposes has a $\frac{1}{n}$ chance of proposing to a given university. If the second student who proposes is rejected by their first choice, it follows that they have a $\frac{1}{n-1}$ chance of selecting one of the remaining $n - 1$ universities in particular. In general, the probability of proposing to a given university is $\frac{1}{n-i}$, where i is the number of universities a student has already proposed to. In the Coupon Collector Problem, each distinct card drawn also changed the probability of the subsequent rounds.

By coupling the probabilities, we observe that having n unique proposals is equivalent to having n unique cards. In the Coupon Collector Problem, the expected number of packs was determined to be $\Theta(n \log n)$. However, the key difference in deferred acceptance is that a given university cannot be proposed to twice by a given student. This means that the probability of an unmatched student proposing to the other universities *increases* during each subsequent round, ultimately removing the lower bound on the expectation of X_i . Therefore, $\mathbb{E}[\sum_i X_i] = O(n \log n)$.