This page is intentionally left blank.
Please do not open this problem sheet until you are told to do so.

4810-1183 Approximation and Online Algorithms with Applications – Final Examination

Problem 1

We discuss randomized rounding in Lecture Note 5. There, you might have a question why we have to set the success probability of each execution to 0.5. In this problem, we will try to set the success probability in a different way.

From the next question, we will try to have the success probability of each execution equal to $1/|E|$.

Question 1.1: When we execute the randomized rounding algorithm for n times, what is the probability of having at least one execution being successful?

Question 1.2: From Question 1.1, discuss why, when $n = 50|E|$, the success probability of the algorithm is as small as $1 - e^{-50}$.

Question 1.3: Calculate the time complexity of the randomized rounding when $n = 50|E|$. Discuss why our algorithm is still efficient there.

Question 1.4: When we set $x_i' = 1$ with probability $1 - (1 - x_i)^t$, calculate the appropriate value of t.

Question 1.5: What is the expected value of SOL in this case?

Question 1.6: Discuss why the approximation result we have in this case is not that better than what we have in the lecture note.

Question 1.7: Discuss why we cannot have an approximation scheme (guess lecture) by just increasing the value of n.

Problem 2

In the problem statement of the k-center problem (Lecture Note 6), we assume triangle inequality, i.e., for all houses i, j, j', we have $d(i, j') \leq d(i, j) + d(j, j')$. We will consider the case that we do not have that assumption in this problem.

Question 2.1: In the lecture note, we assume that there is a library for the 1.9999-approximation k-center problem and use the library to solve the dominating set problem. There, the distance d that we send to the k-center library is as follows:

$$d(i,j) = \begin{cases} 0 & if\ i = j \\ 1 & if\ \{i,j\} \in E \\ 2 & \text{Otherwise} \end{cases}$$

Discuss why this distance satisfies the triangle inequality.

Question 2.2: Now, let us consider the following distance.

$$d(i,j) = \begin{cases} 0 & if\ i = j \\ 1 & if\ \{i,j\} \in E \\ 3 & \text{Otherwise} \end{cases}$$

If we use that in the algorithm for the dominating set problem in lecture note 6, what is the value of $D_{S'}$ when S' is a dominating set? And, what is the value of $D_{S'}$ when S' is not a dominating set?

Question 2.3: Discuss why, if we can solve 2.9999-approximation k-center problem, we can solve the dominating set problem by updating the distance d as in Question 2.2.

Question 2.4: From your answer in Question 2.3, discuss why, if we do not assume triangle inequality, we are unlikely to have a 2.9999-approximation algorithm for the k-center problem.

Question 2.5: Discuss why, if we do not assume triangle inequality, we are unlikely to have an α-approximation algorithm for the k-center problem for any $\alpha > 0$.

Problem 3

In the machine learning online algorithm that we have discussed in Lecture Note 9, we try to minimize the number of incorrect decision. However, each incorrect decision cost us a different amount of money. For example, let us suppose that the stock price falls 300 dollars on day 1 and 1000 dollars on day 3. Incorrect decisions of buying day 1 and day 3 choose be considered differently. We should try to minimize the sum of money lost by incorrect decisions.

Let us suppose that the amount of money lost must always be a positive integer.

Question 3.1: Formulate the online problem for the above setting by modifying the problem stated in Lecture Note 9. What are inputs, outputs, objective function, and constraints?

From the next question, let assume that a lost on a day is at least p and at most P.

Question 3.2: Discuss why we will lose at least pm if we know from the beginning who is the best expert.

Question 3.3: Discuss why we will lose at most $2.41(m + \lg_2 n) \cdot P$ by using the weighted majority voting in lecture note 9.

Question 3.4: What is the competitive ratio of the algorithm for this setting?

Question 3.5: We will modify the algorithm in this question. Instead of decreasing a weight of experts who makes an incorrect suggestion by half, we will reduce their weight by $1/2^p$ when p is the lost on the day he/she makes an incorrect suggestion. Give an example that the lost for the modified algorithm is larger than $2.41(m + \lg_2 n)$.