Problem 2: Communication Conundrum [Email Submission]

By Frederick Qiu

This problem assumes some knowledge on set theoretic notation - a quick guide to everything you need to know is on the last page. Feel free to come to office hours as well if you need clarification on anything!

Alice is stuck on her theoretical computer science problem set and needs a perfect grade to not fail the class. The problem set consists of Lemmas and Theorems, where each Theorem can be proven using some subset of the Lemmas. Each student is privately given a (not necessarily unique) list of Theorems, of which they must prove at least one to get full points on the problem set.

In an act of desperation, Alice asks her friend Bob for help on the problem set. Luckily for Alice, Bob already proved every Lemma. Unluckily for Alice, collaboration on this problem set is against the Honor Code, so Bob will only send the solutions to Lemmas that he won't be submitting for his own problem set (so the plagiarism checkers don't go off).

Given this, Alice wants to figure out if it's possible for both her and Bob to get full points on the problem set (since if it isn't, she'll have to pay off Bob in boba to take an L on his problem set). However, she also wants to minimize the amount of communication between them, lest they be caught by their fellow students and turned in to the Honor committee.

Part 1 (5 points)

We can think of the Lemmas as a universe $L = \{1, 2, \ldots, n\}$ and a Theorem T as the subset of L required to prove it. We say that a collection of Theorems $A_1, A_2, \ldots, A_m$ are *intersecting* if it holds that $A_i \cap A_j^c = \emptyset$ if and only if $i = j$.

Prove that if $A_i \not\subseteq A_j$ for all i, j, then $A_1, A_2, \ldots, A_m$ are intersecting.

Part 2 (5 points)

Using Part 1, give a lower bound $\gamma(n)$ on the largest intersecting collection of Theorems. Your bound may be loose up to a multiplicative factor of n for full points.

Part 3 (5 points)

In the communication game SetDisjointness we have a universe $Z = \{1, 2, \ldots, z\}$. One player is given $X \subseteq Z$, the other player is given $Y \subseteq Z$, and their goal is to determine whether $X \cap Y = \emptyset$. It is known that SetDisjointness requires around z bits of communication to solve.

Suppose Alice and Bob have a protocol for determining whether they can both get perfect scores on the problem set. Show that they can use their protocol to solve SetDisjointness on a universe of size $\gamma(n)$, and hence must use around $\gamma(n)$ bits of communication. If you showed $\gamma(n)$ is very large in Part 2, then you have shown that Alice and Bob are doomed to get caught by the Honor Committee.

Hint: you want to construct a list of Theorems for Alice and a list of Theorems for Bob such that they can both get perfect scores if and only if $X \cap Y \neq \emptyset$.

Background: Why could this be important?

Reading this part isn't necessary to solving the problem, but it'll give you some more insight on why the challenges you solved aren't some random silly problems!

In our introductory computer science courses, we focus a lot on computation, and specifically, how to design efficient algorithms. Some of you might have also taken courses on computational complexity, where we prove (under some reasonable assumptions) that some problems cannot be solved without using a certain amount of computation. These proofs are generally done via *reductions*: if we want to show that solving problem X efficiently is hard, then we first show that an efficient algorithm to solve problem X can be used to efficiently solve problem Y. If we know that problem Y cannot be solved efficiently, then we have shown impossibility of an efficient algorithm to solve problem X.

In this problem, you also did a reduction - you showed that a protocol for determining whether two perfect scores are possible can be used to solve SetDisjointness, and thus, such a protocol needs to use a *lot* of communication (if you got a good lower bound in Part 2).

However, the reduction we did wasn't in terms of computation, it was in terms of communication between Alice and Bob. This is an example of a problem in *communication complexity*, which is concerned with determining how much communication is required to solve problems where multiple parties hold different pieces of information.

Like computation, communication is an important resource to consider when building practical systems. For example, computers can only send so much information to each other over a network, so when designing distributed algorithms, it's important to make sure the communication protocols aren't prohibitively expensive. Communication complexity has many other applications as well, such as in lower bounds for the memory consumption of algorithms which process large data streams.

Set Theory Bootcamp

A *set* is any unordered collection of elements without repetition. In this problem, we will only consider the natural numbers $1, 2, \ldots$ as being elements. The *size* of a set is the number of elements it contains. We denote $\emptyset$ to be the set of size 0.

For a set S and element i, we say that $i \in S$ if S contains i.

For sets S, T, we say that $S \subseteq T$ (S is a *subset* of T) if for all elements $i \in S$, we also have $i \in T$ (in other words, T has everything that S has).

For two sets S, T, we call the set $S \cap T$ their *intersection*. An element $i \in S \cap T$ if and only if $i \in S$ and $i \in T$ (in other words, $S \cap T$ is the overlap of S and T).

Sometimes, we may define some set U to be the *universe*, meaning we restrict attention only to sets which are subsets of U. When this is the case, for any set $S \subseteq U$, we denote S^c to be its *complement*, meaning that for all $i \in U$, $i \in S^c$ if and only if $i \notin S$ (in other words, S^c has everything that S doesn't have, where the universe U is "everything").

Examples

$\{3, 4, 7\}$ is a set. $\{3, 4, 7, 7\}$ is not a set. $\{3, 4, 7\}$ and $\{3, 7, 4\}$ are the same set.

Define the sets $S = \{3, 4, 7\}$, $T = \{3, 4, 5\}$, $R = \{3, 4, 5, 6, 7\}$, and let the universe be $U = \{1, 2, \ldots, 10\}$.

$3 \in S$ and $4 \in S$, but $5 \notin S$.

$S \subseteq R$ and $T \subseteq R$. Also, $R \not\subseteq S$ and $R \not\subseteq T$. Additionally, $S \not\subseteq T$ and $T \not\subseteq S$.

$S \cap T = \{5, 6\}$. $S \cap T = \{3, 4, 7\}$. $S \cap \{1, 2\} = \emptyset$.

$R^c = \{1, 2, 8, 9, 10\}$. $S \cap S^c = \emptyset$.

How to Submit

Email each part separately to coscon.submit@gmail.com, with subject line *Problem2aSubmission*, *Problem2bSubmission*, or *Problem2cSubmission*, respectively. Put the name of all team members in the email body. Provide ample justification for each part. If you must resubmit, *respond to the thread where you sent your original submission; we cannot guarantee that your resubmission will be graded otherwise.*