CS Theory (Spring '25) February 27, 2025

Lecture Note: Communication Protocols

Instructor: *Josh Alman*

0 Logistics

- Next Tuesday (Mar 4) there will be no in-person lecture. Instead, there will be a video lecture.
- Next Thursday (Mar 6) homework 3 will be due at noon.
- On the thursday after next (Mar 13), there will be the midterm exam. The exam will cover materials until next Tuesday (Mar 4), and it will cover everything from class and homeworks.

  You can use results from class and homeworks, but you cannot use without proof results that have not appeared in class or homework problems. For example, "closure of regular languages under intersection" is not proved in class, and has not appeared in homework problems, so you cannot use it without proof, even if you may have used it in your homework solution (and in that case you should already have a good understanding of the proof :).

  You can bring 1 piece of letter-size (or A4-size) paper "cheat sheet" to the exam. You can prepare it in any way you prefer, including handwriting or printing. You can put anything on it, but we suggest you put statements of results from class on it, such as the pumping lemma.
- After the midterm, it will be the spring break!

1 Communication protocols

Once upon a time, there were two persons, Alice and Bob. Suppose Alice and Bob have $x \in \{0,1\}^*$ and $y \in \{0,1\}^*$ as inputs respectively, and their goal is to compute $f(x,y)$ for some function $f : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}$ and for both party to learn $f(x,y)$. To achieve this Goal, Alice and Bob communicates. Their communication might be something like the following:[1]

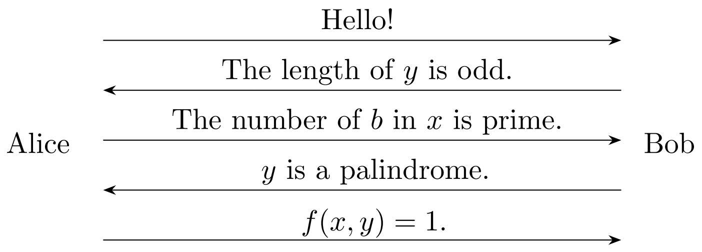

Figure 1: An example of communication between Alice and Bob.

[1] This is just an arbitrary example.

Below are a few examples of communication problems, that is, functions for which we study communication protocols. The convention is to use all capital letters for communication problems.

Example 1. $\text{EQUALITY}(x, y) = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}$.

Example 2. $\text{MAJORITY}(x, y) = \begin{cases} 1 & \textit{if } xy \textit{ contains more 1's than 0's} \\ 0 & \textit{otherwise} \end{cases}$.

Example 3. $\text{PARITY}(x, y) = \begin{cases} 1 & \textit{if } xy \textit{ contains an odd number of } 1\textit{'s} \\ 0 & \textit{otherwise} \end{cases}$.

For any communication problems, the following naïve protocol always works.

Example 4. *Alice first sends her whole input x to Bob. Bob then computes $f(x, y)$ and sends it to Alice. This protocol takes $O(|x|) = O(n)$ bits of communication, where $n := |xy|$.*

However, for many communication problems f, there are much better protocols than the naïve one, in terms of total bits of communication. For example, we will see soon that MAJORITY has a protocol with $O(\log n)$ communication, and PARITY has a protocol with $O(1)$ communication.

In this and the next lecture, we will focus on the following question:

> For a given communication problem, what is the least amount of communication any protocol for the problem have to use?

Definition 5. We call the number of bits of communication used by a protocol the *communication complexity* of the protocol. We also define the communication complexity of a communication problem to be the minimum communication complexity among all protocols for the problem.

The rationale that we focus on communication and not the local computation by Alice or Bob themselves is that communication usually takes much longer than local computation. For example, Alice might be in New York while Bob is in Hawaii. As another example, when designing chips, it is common for the bottleneck to be the communication between different parts of the chip, and this is also one of the reasons why people started to study communication complexity.

Now we discuss better-than-naïve protocols for the MAJORITY and PARITY problems.

Example 6 (Protocol for MAJORITY)**.** *Alice first sends the number of 1's in x and the number of 0's in x to Bob. Bob computes the number of 1's in y and the number of 0's in y, and then computes the total number of 1's in xy and the total number of 0's in xy, using Alice's message. Bob compares the two numbers and sends the answer to Alice.*

The protocol takes $O(\log |x|) = O(\log n)$ communication, where $n = |xy|$, since the two numbers Alice sends to Bob can take value as big as $|x|$.

Example 7 (Protocol for PARITY)**.** *Alice sends to Bob whether the number of 1's in x is odd. Bob computes whether the number of 1's in y is odd, and then computes the answer: 1 if one of the number of 1's in x and the number of 1's in y is odd and the other is even, and 0 is both of these numbers are odd or both are even. Bob then sends the result to Alice.*

There are 2 bits of communication in this protocol. We can prove that this is optimal, as Bob needs 1 bit of message from Alice to learn about the parity of the number of 1's in x, and similarly Alice needs 1 bit of message from Bob.

Remark 8. Bob already knows the answer before sending it to Alice. However, our requirement is that both Alice and Bob know the solution, so Bob has to send it to Alice. Actually, in most protocols, the last message is one party sending the final answer to the other party.

For EQUALITY, the naïve protocol uses $O(n)$ communication, and next time we will see that this is optimal.

One may notice that PARITY and MAJORITY are all defined in the following form: $f(x, y) = 1$ if the string xy satisfies some property, and $f(x, y) = 0$ otherwise. We can generalize this as follows.

Definition 9. Given a language L over $\{0, 1\}$, define the associated communication problem of L to be

$$f_L(x, y) := \begin{cases} 1 & xy \in L \\ 0 & xy \notin L \end{cases}.$$

We can see that PARITY $= f_{L_P}$ where

$$L_P := \{w \in \{0, 1\}^* \mid \text{the number of 1's in } w \text{ is odd}\},$$

and MAJORITY $= f_{L_M}$ where

$$L_M := \{w \in \{0, 1\}^* \mid \text{the number of 1's in } w \text{ is greater than the number of 0's in } w\}.$$

However, not all communication problems are associated with a language. For example, for EQUALITY, one may be tempted to define

$$L_E := \{ww \mid w \in \{0, 1\}^*\}.$$

However, $f_{L_E} \neq$ EQUALITY. As a counterexample, if $x = 000$, $y = 0$, then $f_{L_E}(x, y) = 1$, while EQUALITY$(x, y) = 0$.

We leave a question to think about at the end.

Question 10. *Is it true that f_L has $O(1)$ communication protocol if and only if L is regular?*