CMSC 28100

Introduction to Complexity Theory

Autumn 2025
Instructor: William Hoza

~~Which problems~~

~~can be solved~~

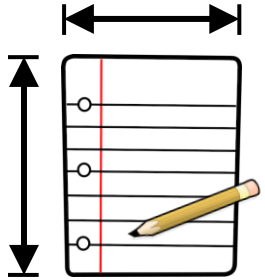
~~through computation?~~

Complexity theory:

The study of computational resources

Computational resources: Fuel for algorithms

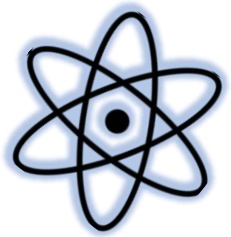
TIME



SPACE

RANDOMNESS

ADVICE



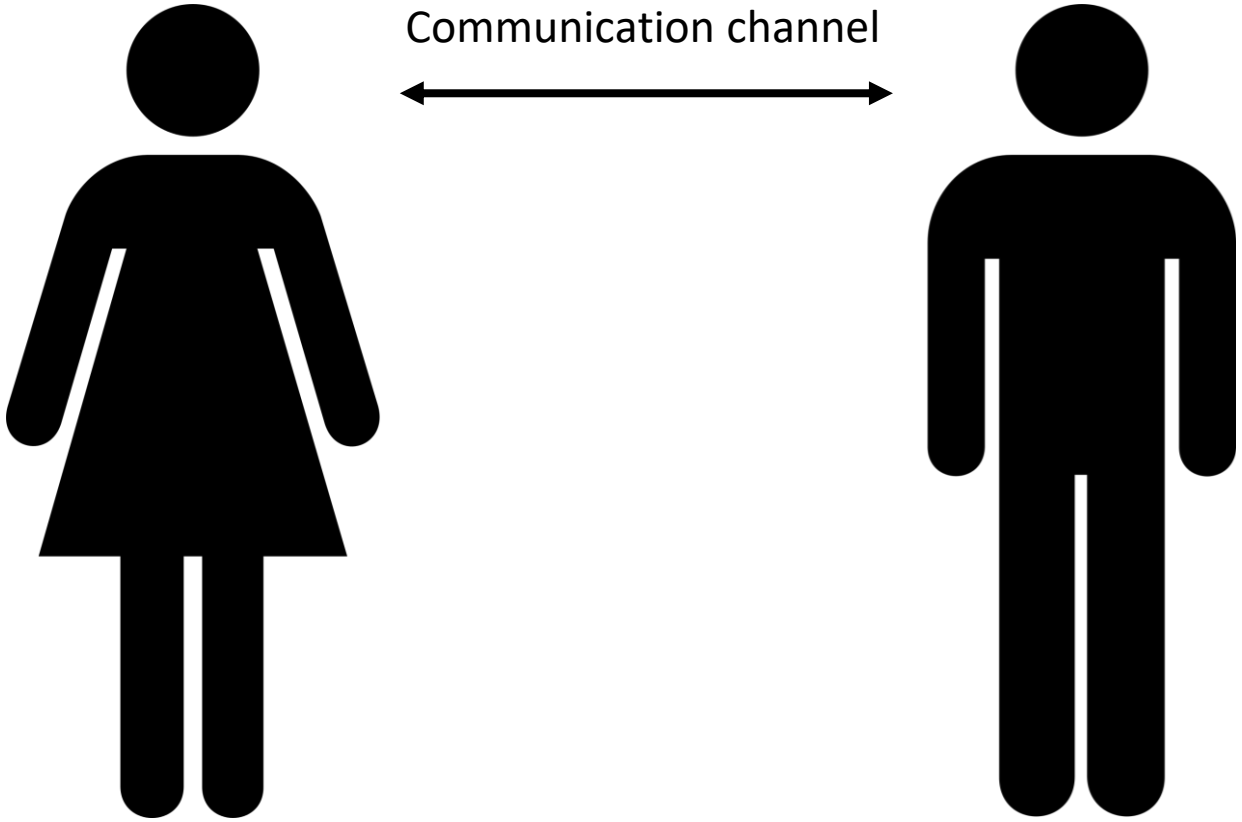
QUANTUM PHYSICS

COMMUNICATION

Communication Complexity

Communication complexity

- Goal: Compute $f(x, y)$ using as little communication as possible
- In each round, one party sends a single bit; the other party listens
- At the end, both parties announce $f(x, y)$



The equality function

- We will focus on the case $f = \mathrm{EQ}_n$
- $\mathrm{EQ}_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$
- Definition:

$$\mathrm{EQ}_n(x,y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}$$

- “Does your copy of the database match my copy?”

Protocols for equality

Protocol A:

1) Alice sends $x \in \{0,1\}^n$
2) Bob sends $\mathrm{EQ}_n(x, y) \in \{0,1\}$

$n + 1$ bits of communication

Protocol B:

1) For $i = 1$ to n:
   a) Alice sends x_i
   b) Bob sends a bit indicating whether $x_i = y_i$

$2n$ bits of communication
(in the worst case)

Communication complexity of equality

- Is there a better protocol?

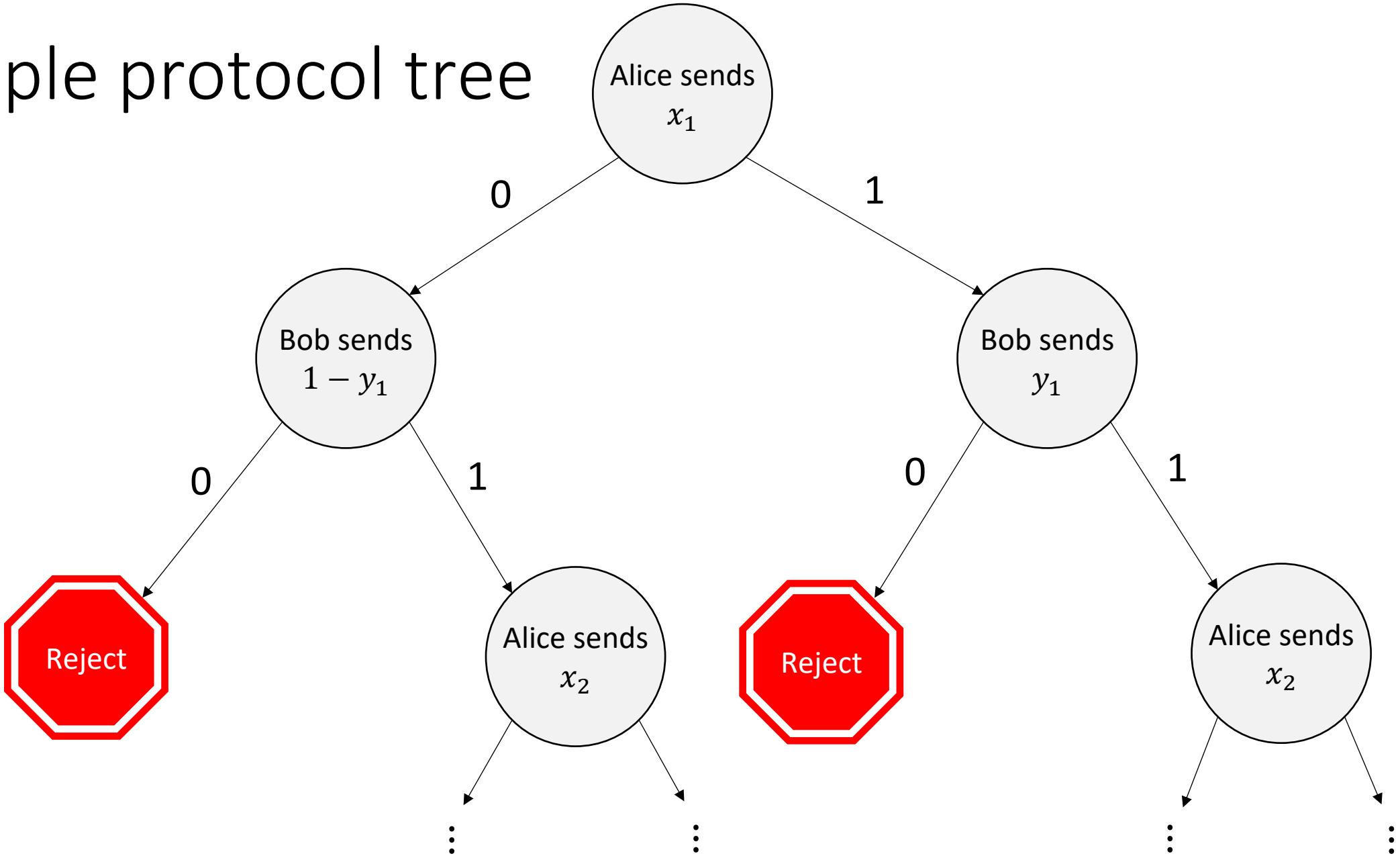
Theorem: Every deterministic communication protocol for EQ_n uses at least $n + 1$ bits of communication in the worst case

- Before we can prove it, we must clarify how we model communication protocols mathematically

Communication protocol model

- Idea: We model a communication protocol as a binary tree
- We start at the root node
- Someone transmits a zero ⇔ We move to the left child
- Someone transmits a one ⇔ We move to the right child
- (Alice and Bob both know where we are in the tree)

Example protocol tree



Rigorously defining communication protocols

- A deterministic communication protocol with n-bit inputs is a rooted binary tree π with the following features
  - The vertex set V is partitioned into $V = V_{\text{Alice}} \cup V_{\text{Bob}} \cup V_{\text{Accept}} \cup V_{\text{Reject}}$
  - Each vertex $v \in V_{\text{Alice}} \cup V_{\text{Bob}}$ has two children (ℓ and r) and is labeled with a function $\delta_v: \{0,1\}^n \to \{\ell, r\}$
  - Each vertex $v \in V_{\text{Accept}} \cup V_{\text{Reject}}$ has zero children

Rigorously defini

- For $x, y \in \{0,1\}^n$, we de
  - Let $v_0 =$ the root vertex
  - If $v_i \in V_{\text{Alice}}$, then let $v_{i+1} = \delta_{v_i}(x)$
  - If $v_i \in V_{\text{Bob}}$, then let $v_{i+1} = \delta_{v_i}(y)$
  - If $v_i \in V_{\text{Accept}} \cup V_{\text{Reject}}$, then let $\text{leaf}(x, y) = v_i$
- We say that π accepts (x, y) if $\text{leaf}(x, y) \in V_{\text{Accept}}$
- We say that π rejects (x, y) if $\text{leaf}(x, y) \in V_{\text{Reject}}$

In this model, what happens if Alice and Bob speak at the same time?

A: Trick question. In this model, they never speak simultaneously

B: Only one of the messages is successfully transmitted

C: Both of the messages are successfully transmitted

D: Neither message is successfully transmitted

Respond at PollEv.com/whoza or text “whoza” to 22333

Communication complexity

- We say that π computes f if for every $x, y \in \{0,1\}^n$,
  - If $f(x,y) = 1$, then π accepts (x,y)
  - If $f(x,y) = 0$, then π rejects (x,y)
- The cost of the communication protocol π is the depth of the tree, i.e., the length of the longest path from the root to the leaf
- (Cost = number of rounds = number of bits of communication)

Leaf structure

- Suppose π is a communication protocol computing EQ_n and $x, y \in \{0,1\}^n$

> **Lemma:** If $\mathrm{leaf}(x,x) = \mathrm{leaf}(y,y)$, then $x = y$

- **Proof (sketch):** Suppose $\mathrm{leaf}(x,x) = \mathrm{leaf}(y,y) = v$ (accepting leaf)
- Let $v_0, v_1, \ldots, v_T$ be the vertices from the root to v
- Then $\delta_{v_i}(x) = \delta_{v_i}(y) = v_{i+1}$ for every i
- Therefore, $\mathrm{leaf}(x,y) = v$, so π accepts (x,y), so $x = y$

Communication complexity of equality

Theorem: Every deterministic communication protocol that computes EQ_n has cost at least $n + 1$

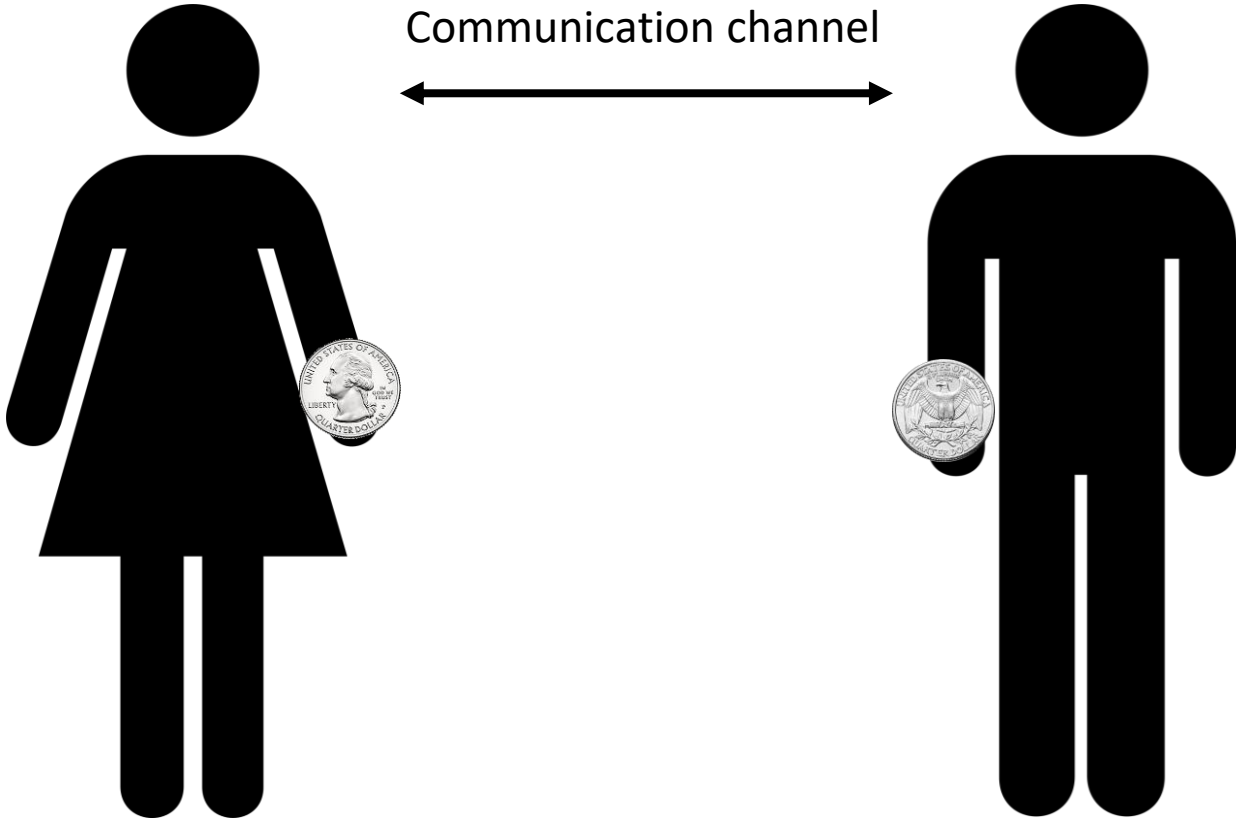
- **Proof:** By the lemma, there are 2^n accepting leaves $\mathrm{leaf}(x, x)$
- There is also at least one rejecting leaf, so there are $> 2^n$ leaves total
- Therefore, there is a leaf at depth $> n$

Communication complexity of EQ_n

- We just proved that computing EQ_n requires $n + 1$ bits of communication
- However, there is a loophole!
- Our impossibility proof only applies to deterministic protocols!

Randomized communication complexity

- In a randomized communication protocol, Alice and Bob are permitted to make decisions based on coin tosses



Randomized communication protocols

- Mathematically, we model a randomized communication protocol with n-bit inputs as a deterministic communication protocol with $(n + r)$-bit inputs for some $r \geq 0$
- Alice holds xu, where $x \in \{0,1\}^n$ and $u \in \{0,1\}^r$
- Bob holds yw, where $y \in \{0,1\}^n$ and $w \in \{0,1\}^r$
- Interpretation: x, y are the "actual inputs," and u, w are the coin tosses

Randomized protocols: Accepting/rejecting

- Experiment: Pick $u, w \in \{0,1\}^r$ independently and uniformly at random
- We say that π accepts (x, y) if π accepts (xu, yw)
- We say that π rejects (x, y) if π rejects (xu, yw)

$$\Pr[\pi \text{ accepts } (x,y)] = \frac{|\{(u,w) : \pi \text{ accepts } (xu, yw)\}|}{2^{2r}}$$

Randomized protocols: Computing a function

- Let $f: \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ and let $\delta \in [0,1]$
- We say that π computes f with error probability δ if for every $x, y \in \{0,1\}^n$:
  - If $f(x,y) = 1$, then $\Pr[\pi \text{ accepts } (x,y)] \geq 1 - \delta$
  - If $f(x,y) = 0$, then $\Pr[\pi \text{ accepts } (x,y)] \leq \delta$

Randomized communication complexity of EQ_n

- Let $\delta > 0$ be any constant

Theorem: For every $n \in \mathbb{N}$, there exists a randomized communication protocol with cost $O(\log n)$ that computes EQ_n with error probability δ

- Randomized protocols are exponentially better than deterministic protocols!
- Proof: Next three slides

Randomized protocol for $\mathbf{EQ}_n$

- Think of $x, y \in \{0,1\}^n$ as numbers $x, y \in \{0, 1, \dots, 2^n - 1\}$
- Let $p_1 \leq p_2 \leq p_3 \leq \cdots$ be the sequence of all prime numbers
- **Protocol:**
  1. Alice picks $i \in \{1, 2, \dots, n/\delta\}$ uniformly at random (WLOG, n/δ is a power of two)
  2. Alice sends i and $x \bmod p_i$
  3. Bob sends a bit indicating whether $x \bmod p_i = y \bmod p_i$
  4. If so, they accept, otherwise, they reject

Analysis of the protocol: Correctness

Protocol:

1. Pick $i \in \{1, 2, \dots, n/\delta\}$ u.a.r.
2. Send i and $x \bmod p_i$
3. Check whether $x \equiv y \bmod p_i$

- If $x = y$, then $\Pr[\text{accept}] = \Pr[x \equiv y \bmod p_i] = 1$ ✔
- If $x \neq y$, then $\Pr[\text{accept}] = \Pr[x \equiv y \bmod p_i] = \Pr[p_i \text{ divides } |x - y|]$
- Let BAD be the set of prime numbers that divide $|x - y|$
- $2^{|\text{BAD}|} \leq \prod_{p \in \text{BAD}} p \leq |x - y| < 2^n$
- $\Pr[\text{accept}] = \Pr[p_i \in \text{BAD}] \leq \frac{|\text{BAD}|}{n/\delta} < \frac{n}{n/\delta} = \delta$ ✔

Analysis of the protocol: Efficiency

Protocol:

1. Pick $i \in \{1, 2, \dots, n/\delta\}$ u.a.r.
2. Send i and $x \bmod p_i$
3. Check whether $x \equiv y \bmod p_i$

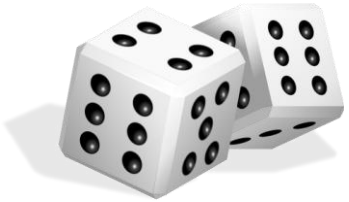
- Sending i costs $O(\log n)$ bits of communication ✔
- Sending $x \bmod p_i$ costs $O(\log p_i)$ bits of communication
- How big is p_i (the i-th prime)?

Chebyshev's Estimate: Let p_k be the k-th prime. Then $p_k = O(k \cdot \log k)$.

  - (Proof omitted)

- Therefore, $\log p_i = \log\big(O(n \cdot \log n)\big) = \log\big(o(n^2)\big) = O(\log n)$ ✔

Recap: The power of randomness



- Is randomness useful?
- For communication protocols: Yes!
- For Turing machines: Probably not much
  - Conjecture: $\mathrm{P} = \mathrm{BPP}$