CMSC 28100

Introduction to Complexity Theory

Spring 2025
Instructor: William Hoza

Midterm exam

- Midterm exam will be in class on **Wednesday, April 23**
- To prepare for the midterm, you only need to study the material prior to this point
- The midterm will be about decidability, undecidability, time complexity, and P

Robustness of P, revisited

- Let $Y \subseteq \{0, 1\}^*$. If $Y \notin \mathrm{P}$, then Y cannot be decided by…
  - A poly-time one-tape Turing machine
  - A poly-time multi-tape Turing machine
  - A poly-time word RAM program
- **OBJECTION:** "This still leaves open the possibility that I could somehow build a device that decides Y in polynomial time."

Extended Church-Turing Thesis

~~**Extended Church-Turing Thesis:**~~

~~For every $Y \subseteq \{0,1\}^*$, it is physically possible to build a device that decides Y in polynomial time if and only if $Y \in \mathrm{P}$.~~

- If it were true, the thesis would justify studying P
- But the thesis is probably false!
- Two key challenges: Randomized Computation and Quantum Computation

Randomized computation

- Researchers often use randomness to answer questions
  - Random sampling for opinion polls
  - Randomized controlled trials in science/medicine
- What if we incorporate this ability into our computational model?

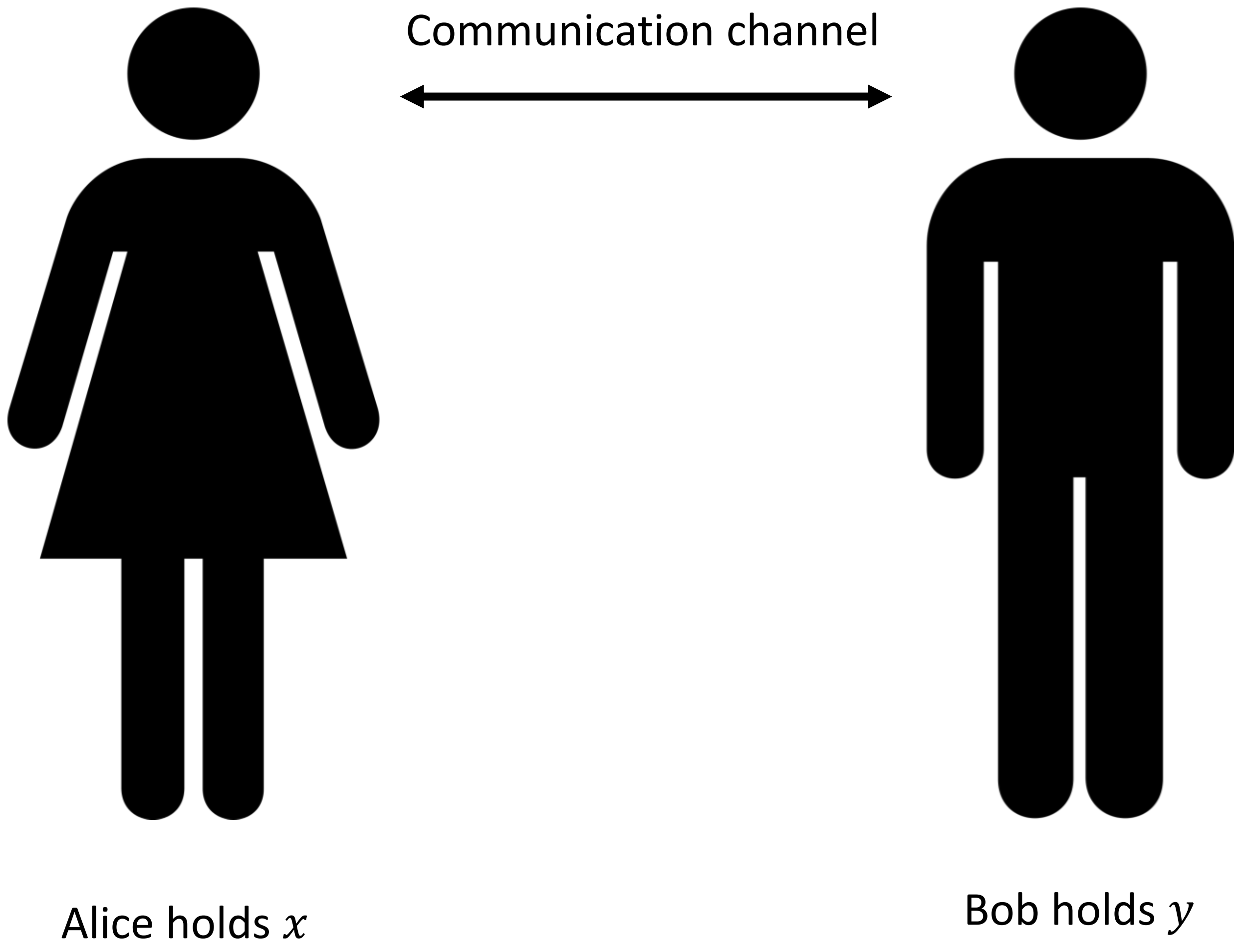
Randomized computation

- Eventually, we will define and study randomized Turing machines
- First, to build intuition, let's study the role of randomness in a different situation

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 $f = \mathrm{EQ}_n$
- $\mathrm{EQ}_n \colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$
- 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?

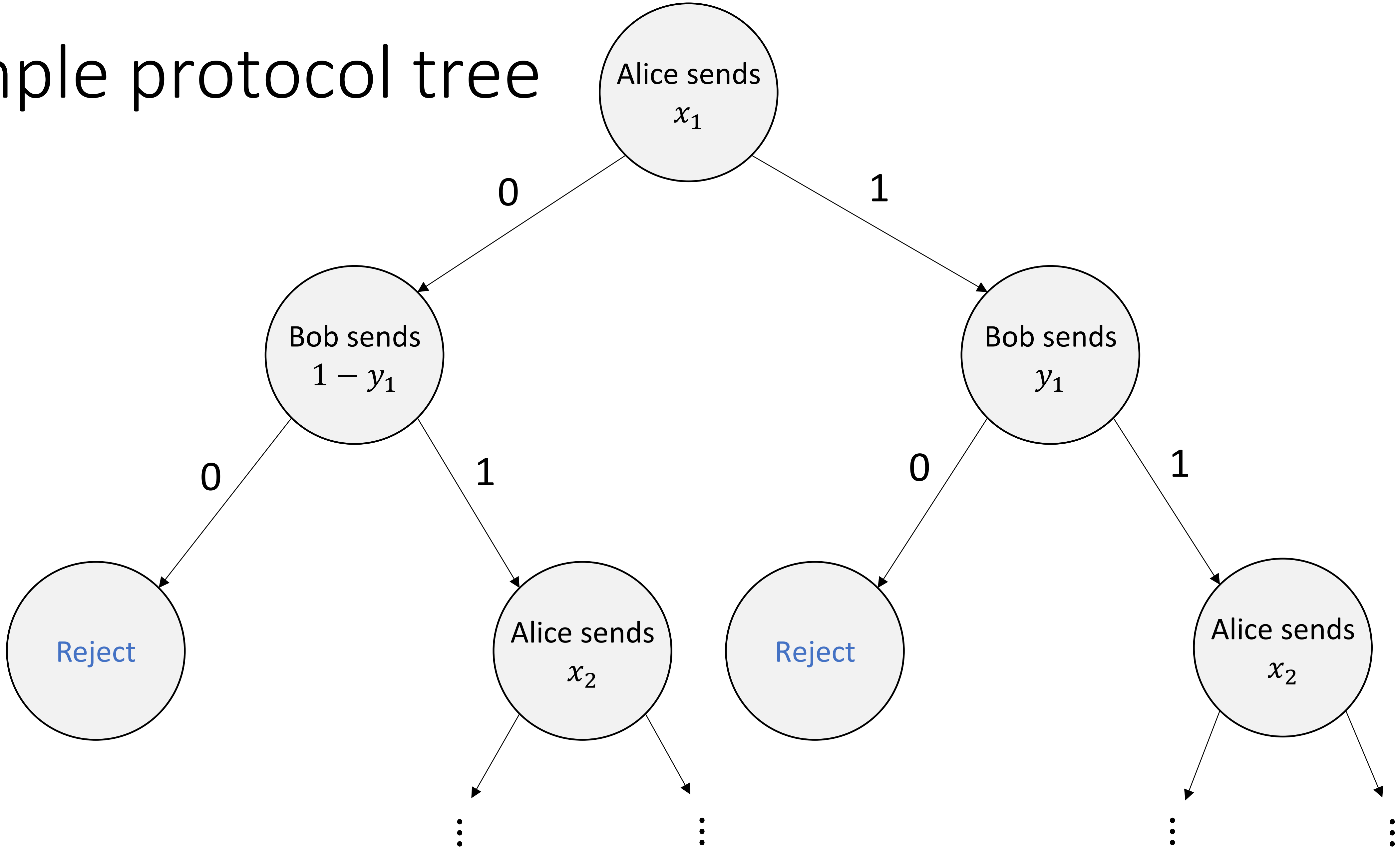
> **Theorem:** Every deterministic communication protocol for EQ_n uses at least $n + 1$ bits of communication in the worst case

- 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 $\delta_v \colon \{0,1\}^n \to \{\ell, r\}$
  - Each vertex $v \in V_{\text{Accept}} \cup V_{\text{Reject}}$ has zero children

Rigorously defini

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

- 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 $\text{leaf}(x, y) \in V_{\text{Reject}}$

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)

Rectangle lemma

$$\begin{tikzcd}
(x,y) \arrow[r, dash, dashed] \arrow[d, dash, dashed] & (x',y) \arrow[d, dash, dashed] \\
(x,y') \arrow[r, dash, dashed] & (x',y')
\end{tikzcd}$$

- Let π be any communication protocol with n-bit inputs
- Let $x, x', y, y' \in \{0,1\}^n$ and let v be any leaf

Rectangle Lemma: If $\mathrm{leaf}(x,y) = \mathrm{leaf}(x',y') = v$,

then $\mathrm{leaf}(x,y') = \mathrm{leaf}(x',y) = v$

- **Proof (sketch):** Let $v_0, v_1, \dots, v_T$ be the vertices from the root to v
- If $v_i \in V_A$, we must have $\delta_{v_i}(x) = \delta_{v_i}(x') = v_{i+1}$. Similarly if $v_i \in V_B$

Communication complexity of equality

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

- **Proof:** Let π be any communication protocol that computes EQ_n
- Assume WLOG that every leaf is at the same depth m
- Our job is to prove that $m \geq n + 1$

Communication complexity of $\mathbf{EQ}_n$

$$\begin{tikzcd}
(x,x) \arrow[r, dash, dashed] \arrow[d, dash, dashed] & (x,y) \arrow[d, dash, dashed] \\
(y,x) \arrow[r, dash, dashed] & (y,y)
\end{tikzcd}$$

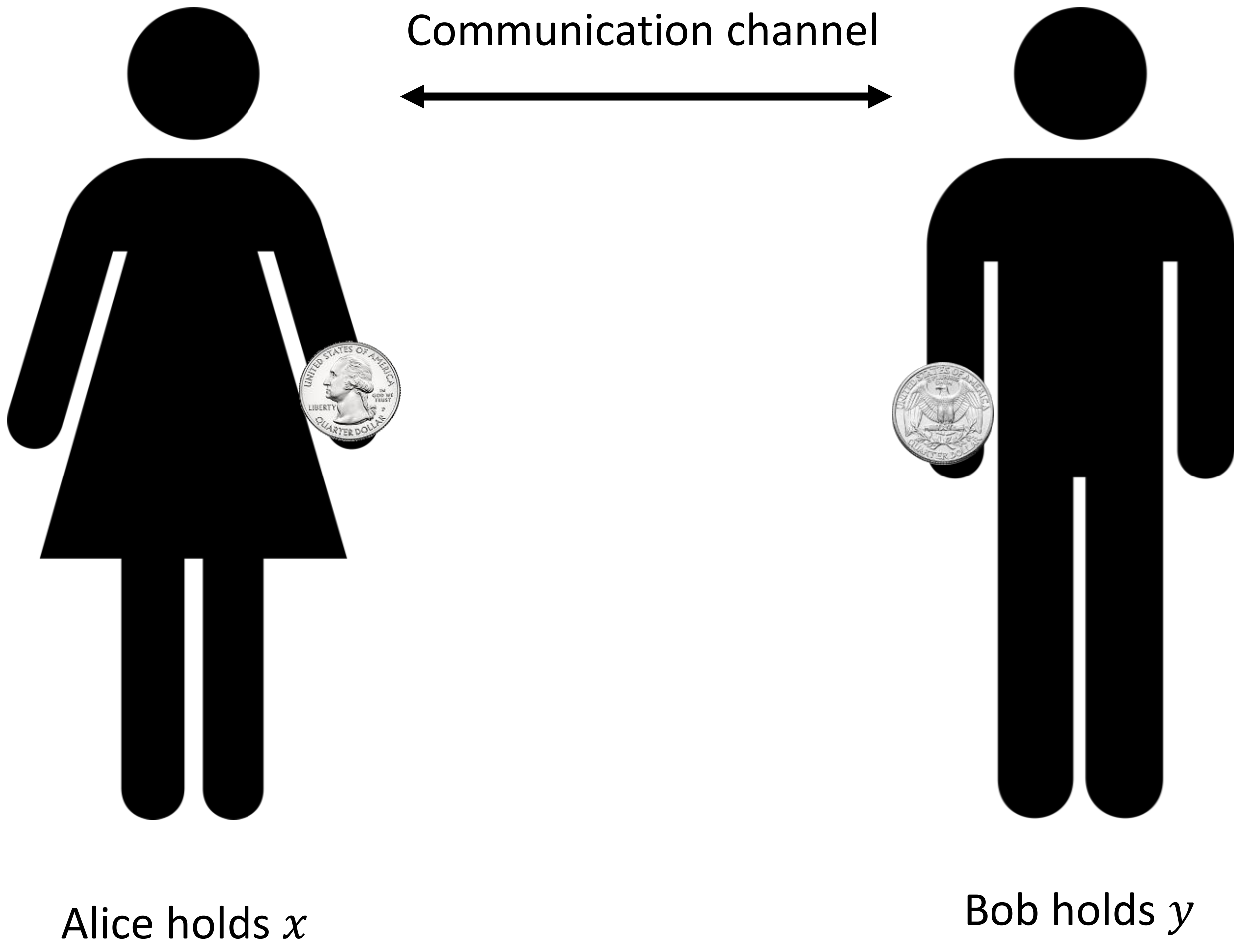
- If $x \neq y$, then $\mathrm{leaf}(x,x) \neq \mathrm{leaf}(x,y)$
- By the rectangle lemma, it follows that $\mathrm{leaf}(x,x) \neq \mathrm{leaf}(y,y)$
- Therefore, $|V_{\mathrm{Accept}}| \geq 2^n$
- Meanwhile, $|V_{\mathrm{Reject}}| \geq 1$
- Therefore, $2^m = |V_{\mathrm{Accept}} \cup V_{\mathrm{Reject}}| > 2^n$, hence $m \geq n+1$

Communication complexity of $\mathbf{EQ}_n$

- 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 $\mathbf{EQ}_n$

- 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!