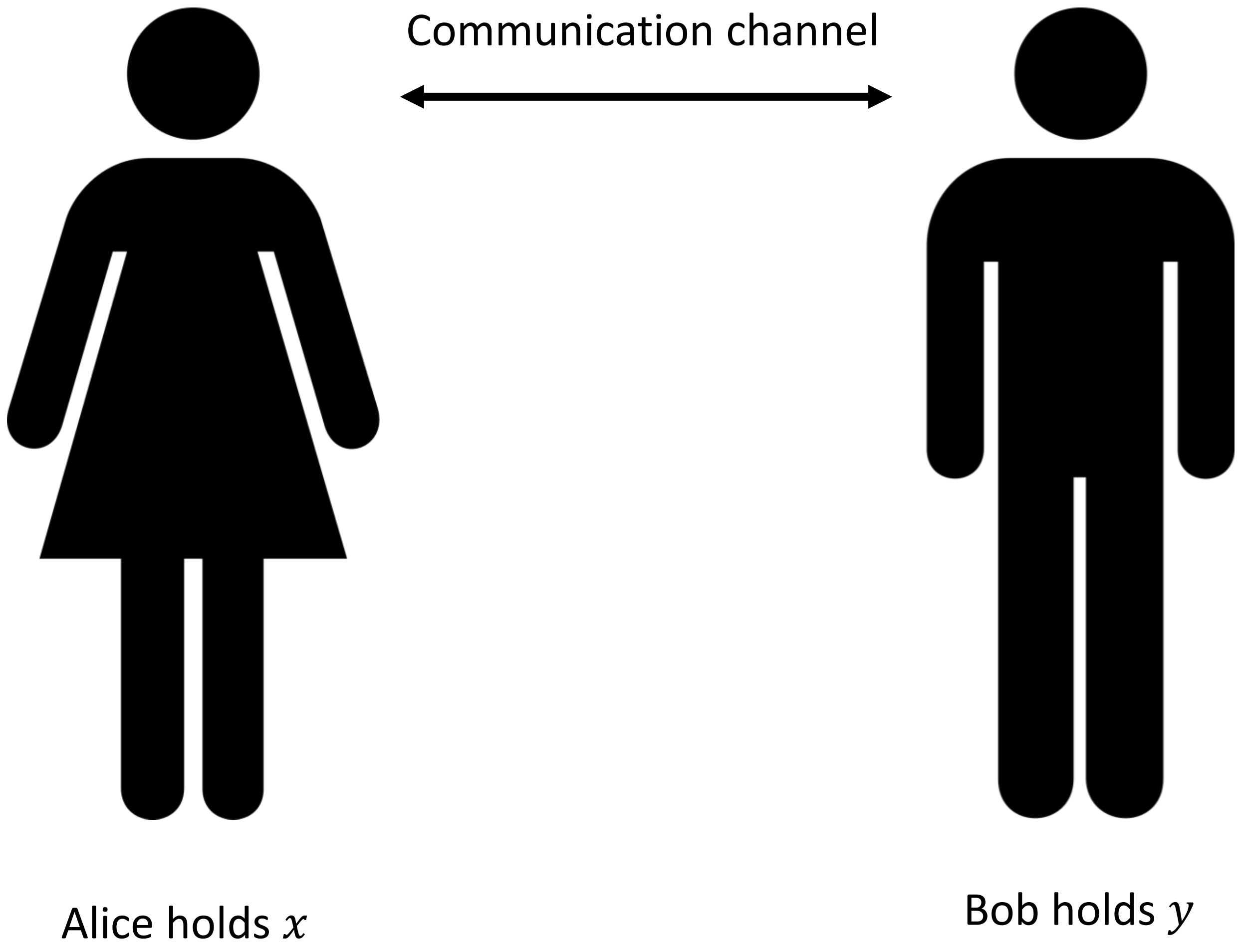
CMSC 28100

Introduction to Complexity Theory

Spring 2024
Instructor: William Hoza

Communication complexity

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



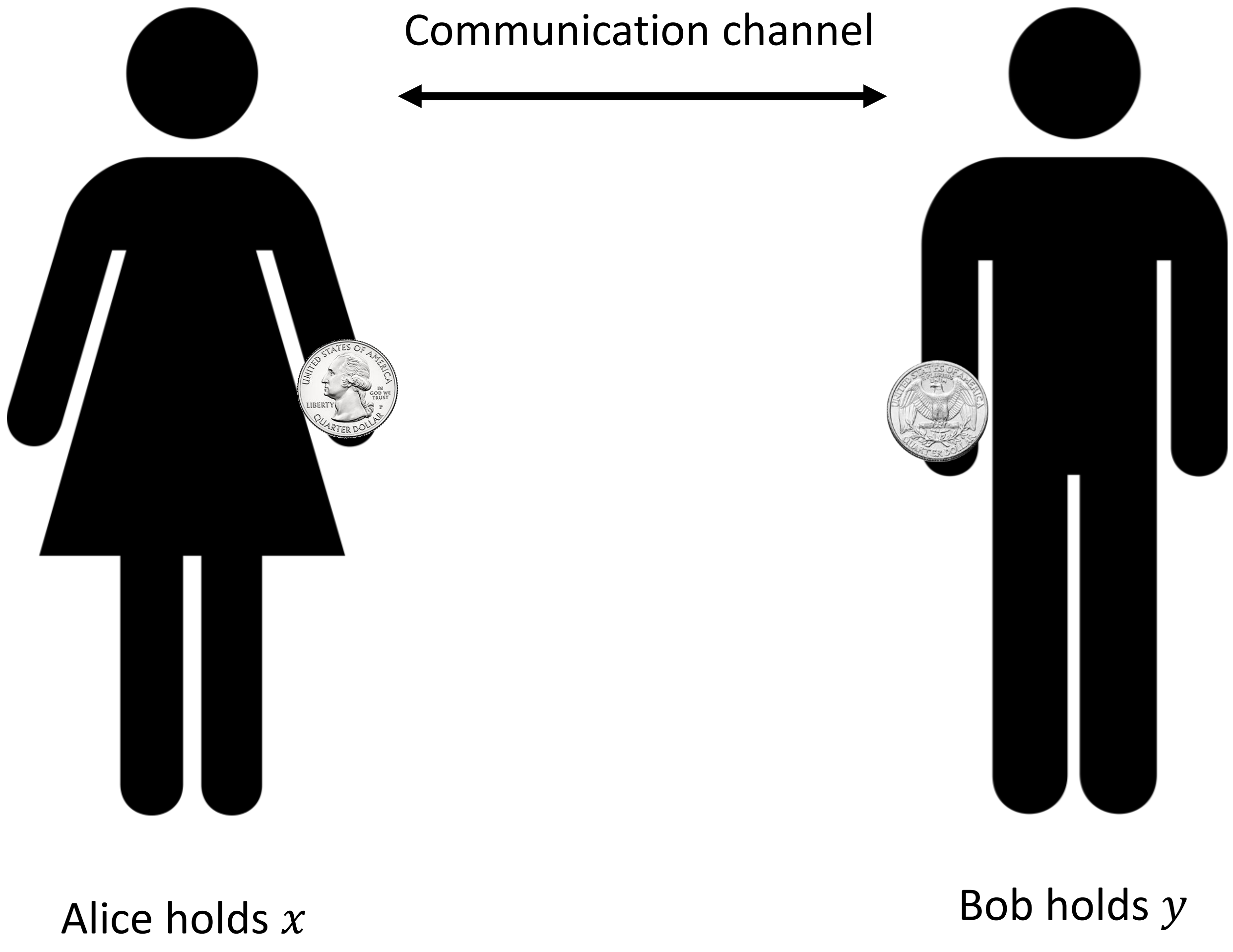
Communication complexity of equality

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

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

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

The output of a randomized protocol

- For each $x, y \in \{0,1\}^n$ and $u, w \in \{0,1\}^r$, we have $\pi(xu, yw) \in \{0,1\}$
- We define $\pi(x, y)$ as follows: Pick $(u, w) \in \{0,1\}^r \times \{0,1\}^r$ uniformly at random, then set $\pi(x, y) = \pi(xu, yw)$
- $\pi(x, y)$ is a random variable. For each $b \in \{0,1\}$, we have

$$\Pr[\pi(x, y) = b] = \frac{|\{(u, w) : \pi(xu, yw) = b\}|}{|\{0,1\}^r \times \{0,1\}^r|}$$

Computing a function with a randomized protocol

- Let $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ and let $\delta > 0$
- Suppose that for every $x, y \in \{0,1\}^n$, we have

$$\Pr[\pi(x,y) = f(x,y)] \geq 1 - \delta$$

- In this case, we say that π computes f with error probability δ

Randomized comm

Which of the following is an accurate description of the protocol?

A: The protocol succeeds on most pairs of inputs

B: The amount of communication is rarely more than $O(\log n)$

C: For every pair of inputs, the protocol is likely to succeed

D: It is likely that for every pair of inputs, the protocol succeeds

Respond at PollEv.com/whoza or text "whoza" to 22333

- 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!

Randomized protocol for $\mathbf{EQ}_n$

- Assume without loss of generality that n/δ is a power of two
- Think of $x, y \in \{0,1\}^n$ as numbers $x, y \in \{0, 1, \dots, 2^n - 1\}$
- Let $p_1, p_2, p_3, \dots$ be the sequence of all prime numbers (in order)
- **Protocol:**
  1. Alice picks $i \in \{1, 2, \dots, n/\delta\}$ uniformly at random and sends $\langle i, x \bmod p_i \rangle$ to Bob
  2. Bob sends a bit indicating whether $x \bmod p_i = y \bmod p_i$
  3. If so, they accept, otherwise, they reject

Analysis of the protocol: Correctness

- If $x = y$, then $\Pr[x \bmod p_i = y \bmod p_i] = 1$ ✓ Now suppose $x \neq y$
- $\Pr[\text{error}] = \Pr[x \bmod p_i = y \bmod p_i] = \Pr[p_i \text{ divides } |x - y|]$
- Let BAD be the set of prime numbers that divide $|x - y|$
- Every prime number is at least 2, so $|\text{BAD}| \leq \log\ |x - y| < n$
- There are n/δ candidate i values, so $\Pr[p_i \in \text{BAD}] \leq \frac{n}{n/\delta} = \delta$ ✓

Analysis of the protocol: Efficiency

- We chose $i \le n/\delta$, so sending i costs $O(\log n)$ bits of communication ✓
- Sending $x \bmod p_i$ costs $O(\log p_i)$ bits of communication
- How big could p_i be?

Theorem: Let p_k be the k-th prime. Then $p_k = \Theta(k \cdot \log k) = o(k^2)$.

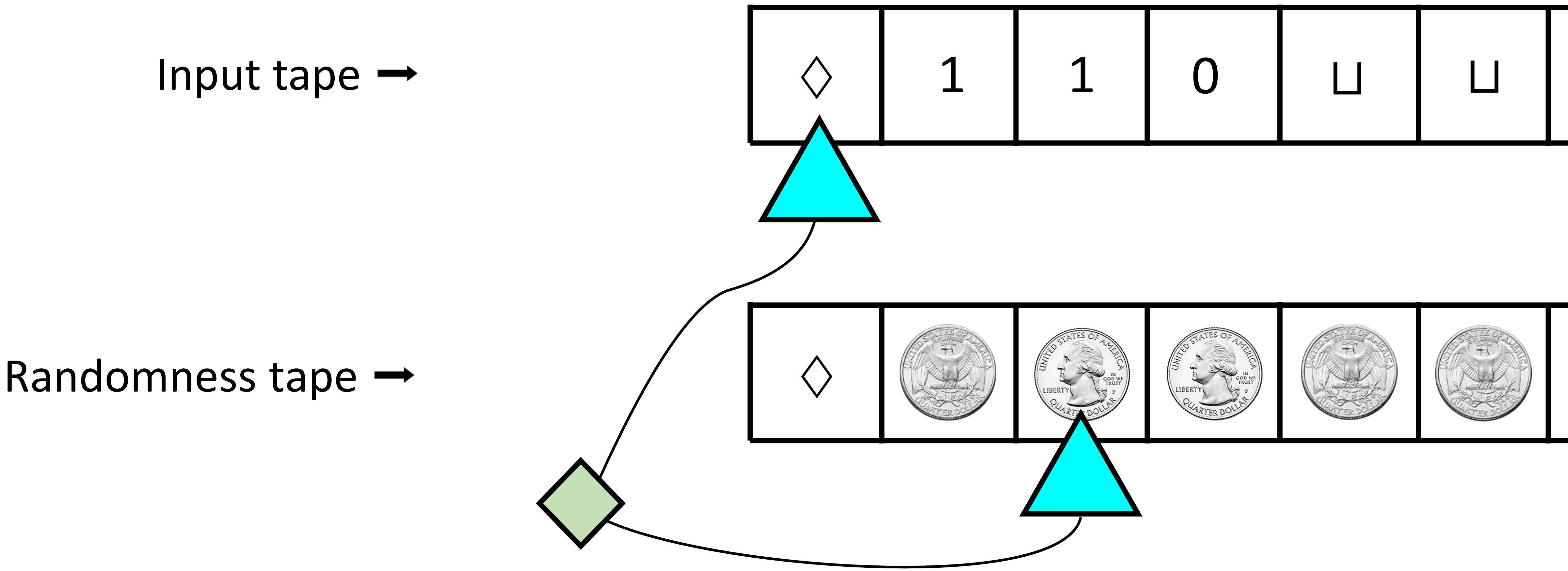
- (Proof is outside the scope of this course)
- We chose $i \le n/\delta$, so $\log p_i = \log\left(o(n^2)\right) = O(\log n)$ ✓

Which problems

can be solved

through computation?

Randomized Turing machines



Randomized Turing machines

- Let $T : \mathbb{N} \to \mathbb{N}$ be a function (time bound)
- **Definition:** A randomized time-T Turing machine is a two-tape Turing machine M such that for every $n \in \mathbb{N}$, every $w \in \Sigma^n$, and every $u \in \{0, 1\}^{T(n)}$, if we initialize M with w on tape 1 and u on tape 2, then it halts within $T(n)$ steps
- Interpretation: w is the input and u is the coin tosses
- (Giving M more than $T(n)$ random bits would be pointless)

Acceptance probability

- Let M be a randomized Turing machine and let $w \in \Sigma^*$
- To run M on w, we select $u \in \{0,1\}^{T(n)}$ uniformly at random and initialize M with w on tape 1 and u on tape 2
- M might accept w sometimes and reject w other times

$$\Pr[M \text{ accepts } w] = \frac{|\{u : M \text{ accepts } w \text{ when tape 2 is initialized with } u\}|}{\left|\{0,1\}^{T(n)}\right|}$$

The complexity class BPP

- Let $L \subseteq \Sigma^*$ be a language
- **Definition:** $L \in \mathrm{BPP}$ if there exists a randomized polynomial-time Turing machine M such that for every $w \in \Sigma^*$:
  - If $w \in L$, then $\Pr[M \text{ accepts } w] \geq 2/3$
  - If $w \notin L$, then $\Pr[M \text{ accepts } w] \leq 1/3$
- "Bounded-error Probabilistic Polynomial-time"