CMSC 28100

Introduction to Complexity Theory

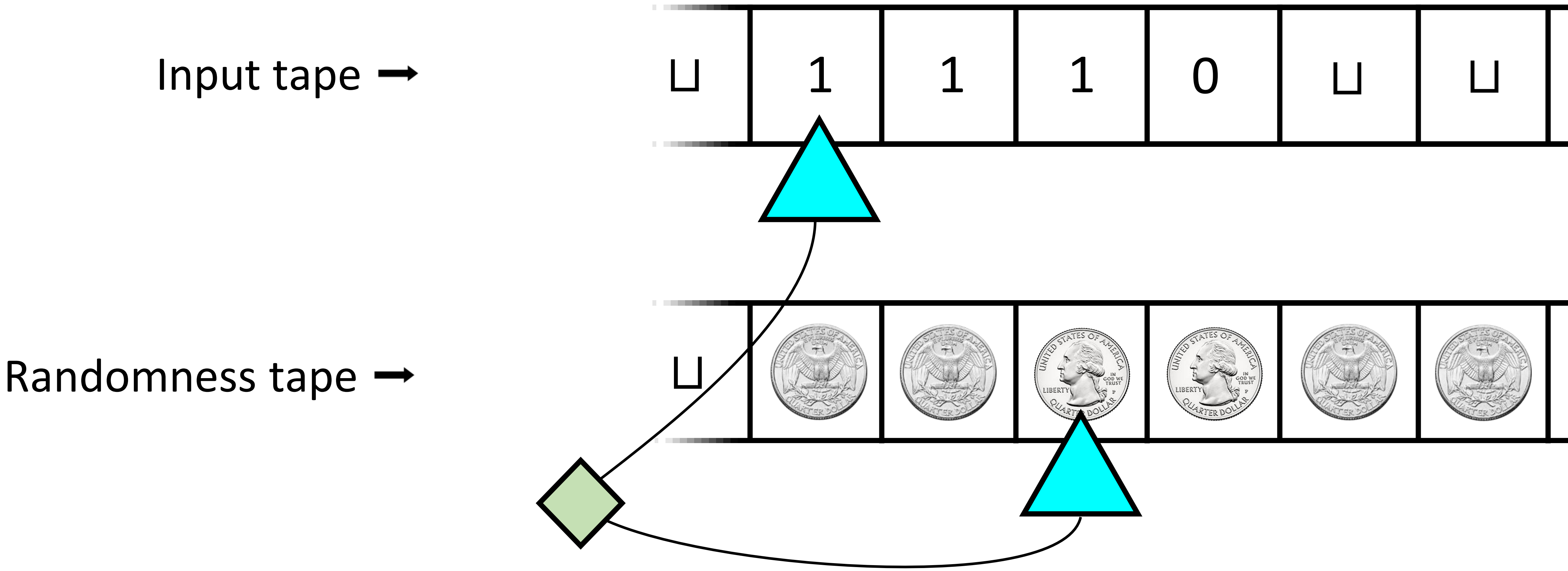
Autumn 2025
Instructor: William Hoza

Which problems

can be solved

through computation?

Randomized Turing machines



The complexity class BPP

- **Definition:** BPP is the set of languages $Y \subseteq \{0,1\}^*$ such that there exists a randomized polynomial-time Turing machine that decides Y with error probability $1/3$
- "Bounded-error Probabilistic Polynomial-time"

Example: High school algebra

- “Expand and simplify: $\underbrace{(x + 1) \cdot (x - 1)}_{\text{This type of expression is called an arithmetic formula}}$”

This type of expression is

called an arithmetic formula

- How difficult is this type of exercise?

Identity testing

- **Problem:** Given an arithmetic formula F, determine whether $F \equiv 0$
- **As a language:**

$$\text{IDENTICALLY-ZERO} = \{\langle F \rangle : F \text{ is an arithmetic formula and } F \equiv 0\}$$

Identity testing example

- Given: $F = (ab + a - b - 1) \cdot (cd - ad + a - c) \cdot (b - e) + (bd + d - b - 1) \cdot (bc + ea - ab - ce) \cdot (1 - a)$
- Expand:

$$F \equiv ab^2cd - eabcd - a^2b^2d + ea^2bd - ab^2c + eabc + a^2b^2 - ea^2b + acdb - eacd - a^2db + ea^2d - acb$$
$$+ eac + a^2b - ea^2 - b^2cd + ebcd + b^2da - ebda + b^2cb - ebc - b^2a + eba - cdb + ecd + dab - eda + cb$$
$$- ec - ab + ea - ea^2bd + eabd + ea^2b - eab - ea^2d + ead + ea^2 - ea + a^2b^2d - ab^2d - a^2b^2 + ab^2$$
$$+ a^2db - adb - a^2b + ab - b^2cda + b^2cd + bcdea - bcde + b^2ca - b^2c - bcea + bce - cdab + cdb + cab$$
$$- cb + cdea - cde - cea + ce$$

- Everything cancels out: $F \equiv 0$

Complexity of identity testing

- Expanding F takes $2^{\Omega(n)}$ time in some cases 😟
- E.g., $F = (x + y) \cdot (x + y) \cdot (x + y) \cdots (x + y)$
- **Open Question:** Is $\text{IDENTICALLY-ZERO} \in \text{P}$?
- Next 5 slides: We will prove $\text{IDENTICALLY-ZERO} \in \text{BPP}$

Identity testing algorithm: Approach

- **Goal:** Figure out whether $F \equiv 0$, where F is an arithmetic formula
- **Strategy:** Compute $F(\vec{x})$ for some $\vec{x}$
- **Rationale:** If $F \equiv 0$, then $F(\vec{x}) = 0$ for all $\vec{x}$ 🙂
- **Difficulty:** Even if $F \not\equiv 0$, there still might be $\vec{x}$ such that $F(\vec{x}) = 0$ ☹️
- How often can this occur?

Counting ro

How many roots can a nonzero degree-d two-variable polynomial have?

A: Up to d

B: Up to d^2

C: It might have infinitely many

D: Only finitely many, but there is no bound in terms of d

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

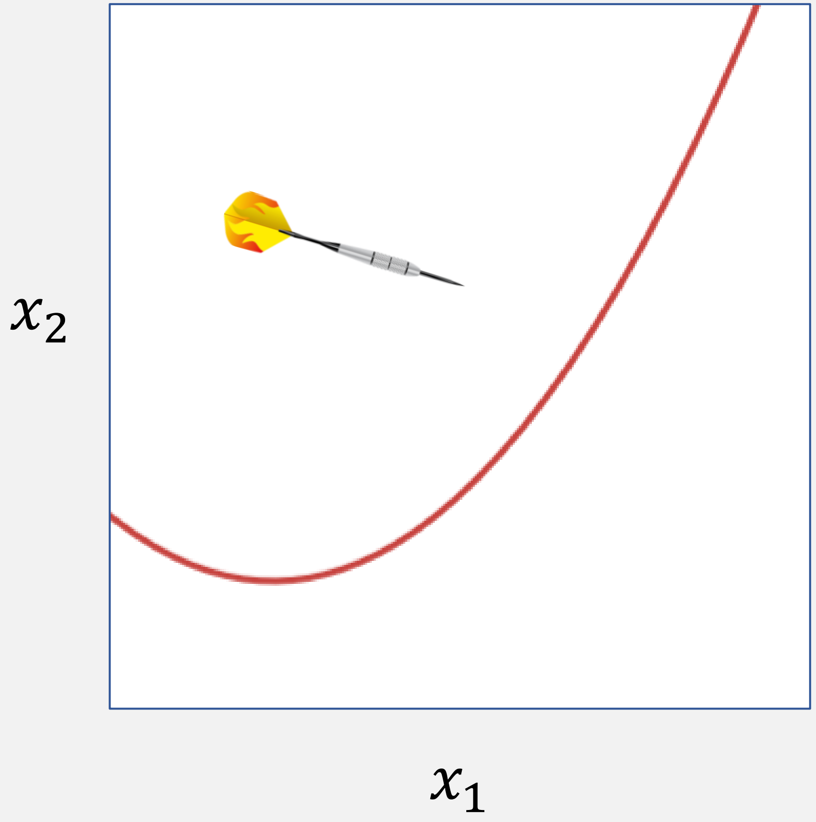
- **Fundamental Theorem of Algebra** $\Rightarrow$ Every nonzero degree-d univariate polynomial has at most d real roots
- What about a multivariate polynomial?

How common are roots?

- Even if $F \not\equiv 0$, it might have infinitely many roots
- Insight: Roots are nevertheless "rare"
- If we pick $\vec{x}$ at random, it is unlikely that $F(\vec{x}) = 0$

Roots of F, where

$$F(\vec{x}) = x_2 - x_1^2$$



Polynomial Identity Lemma

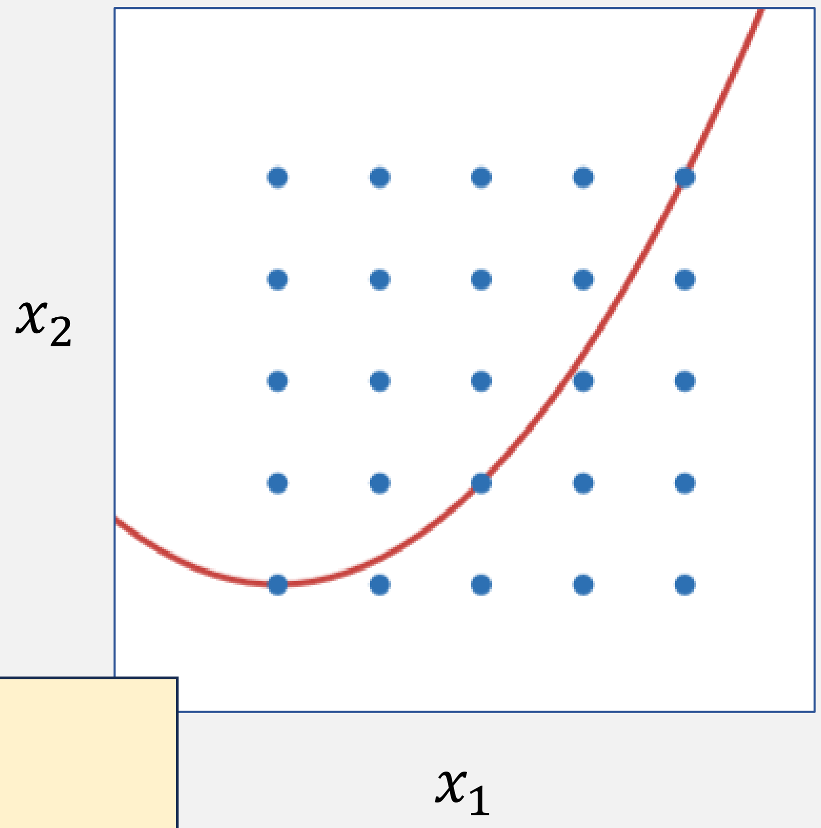
- Let $F : \mathbb{R}^k \to \mathbb{R}$ be a multivariate polynomial of degree at most d in each variable individually
- Let S be a finite subset of $\mathbb{R}$

Polynomial Identity Lemma:

If $F \not\equiv 0$, then $\left|\{\vec{x} \in S^k : F(\vec{x}) = 0\}\right| \leq dk \cdot |S|^{k-1}$

Roots of F, where

$F(\vec{x}) = x_2 - x_1^2$



Proof: On chalkboard

Theorem: $\text{IDENTICALLY-ZERO} \in \text{BPP}$

Given F with k variables and d leaves:

1. Let $S = \{1, \dots, 3dk\}$
2. Pick $\vec{c} \in S^k$ uniformly at random
3. Construct F' by replacing x_i with c_i
4. If $\langle F' \rangle \in \text{EQUALS-ZERO}$, accept, otherwise reject

- Polynomial time ✔
- **Correctness proof**:
- Degree $\leq d$ (can prove by induction)
- If $F \equiv 0$, then P
- If $F \not\equiv 0$, then b

$\Pr[\text{accept}] =$ ... $\frac{\ldots k}{\ldots dk} = \frac{1}{3}$

Which of the following best describes the algorithm?

A: The algorithm behaves correctly on most inputs

B: The amount of time it uses is rarely more than polynomial

C: For every input, the algorithm is likely to behave correctly

D: It is likely that for every input, the algorithm behaves correctly

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

Identity testing: Recap

- We proved $\mathrm{IDENTICALLY\text{-}ZERO} \in \mathrm{BPP}$
- Therefore, we should consider $\mathrm{IDENTICALLY\text{-}ZERO}$ to be tractable
- Does this mean P is a bad model of tractability?
- Not necessarily. Maybe $\mathrm{IDENTICALLY\text{-}ZERO} \in \mathrm{P}$