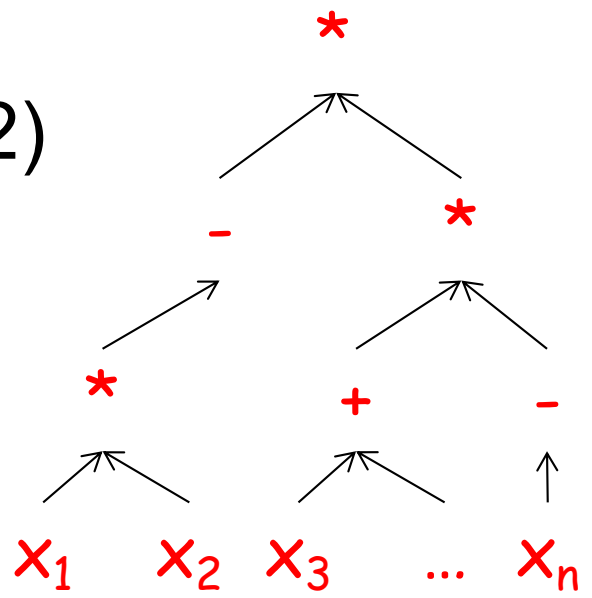
Umans Complexity Theory Lectures

Randomization in

Polynomial Identity Testing + Schwartz-Zippel Lemma

Randomized Algorithms for Polynomial Identity Testing

- Given: polynomial $p(x_1, x_2, \ldots, x_n)$ as arithmetic formula (fan-out 1):
  - multiplication (fan-in 2)
  - addition (fan-in 2)
  - negation (fan-in 1)



Randomized Algorithms for Polynomial Identity Testing

- Let **F** be a field (is closed over addition and multiplication and has both additive and multiplicative inverses).
  - for example $Z_p = \{0,\ldots,p-1\}$ for a prime p
- Question: **Is p identically zero?**
  - i.e., is $p(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbf{F}^n$
  - (assume $|\mathbf{F}|$ larger than degree of $p(\mathbf{x})$)

- "polynomial identity testing" because given two polynomials p, q, we can check the identity $p \equiv q$ by checking if $(p - q) \equiv 0$

Randomized Algorithms for Polynomial Identity Testing

- Try all $|\mathbf{F}|^n$ inputs?
  - may be exponentially many
- multiply out symbolically, check that all coefficients are zero?
  - may be exponentially many coefficients

- **Can randomness help?**
  - i.e., flip coins, allow small probability of wrong answer

Randomized Algorithms for Polynomial Identity Testing

Lemma (Schwartz-Zippel): Let $p(x_1, x_2, \ldots, x_n)$ be a multivariate polynomial over a field **F** with n variables and total degree d.

Let S be any subset of **F**.

Then if p is not identically 0, for uniform random choice of $r_1, r_2, \ldots, r_n$ in S,

$$\Pr_{r_1,r_2,\ldots,r_n \in S}[\, p(r_1, r_2, \ldots, r_n) = 0] \leq d/|S|.$$

(Note: this probability bound does not depend on n.)

Randomized Algorithms for Polynomial Identity Testing

- Proof:
  - induction on number of variables n
  - base case: n = 1, p is univariate polynomial of degree at most d
  - By the fundamental Theorem of Algebra, p has at most d roots, so

$$\Pr[\, p(r_1) = 0] \leq d/|S|$$

Randomized Algorithms for Polynomial Identity Testing

– write $p(x_1, x_2, \ldots, x_n)$ as

$$p(x_1, x_2, \ldots, x_n) = \Sigma_i (x_1)^i p_i(x_2, \ldots, x_n)$$

– Let k = max i for which $p_i(x_2, \ldots, x_n)$ is not identically zero,

– So $p_k(x_2, \ldots, x_n)$ has degree d-k.

– By induction hypothesis:

$$\Pr[\, p_k(r_2, \ldots, r_n) = 0] \leq (d\text{-}k)/|S|$$

– Whenever $p_k(r_2, \ldots, r_n) \neq 0$, $p(x_1, r_2, \ldots, r_n)$ is a univariate polynomial of degree k, so by induction hypothesis:

$$\Pr[p(r_1, r_2, \ldots, r_n) = 0 \mid p_k(r_2, \ldots, r_n) \neq 0] \leq k/|S|$$

Randomized Algorithms for Polynomial Identity Testing

So we have shown both:

(1) $\Pr[\ p_k(r_2, \ldots, r_n) = 0] \leq (d\text{-}k)/|S|$, and

(2) $\Pr[p(r_1,r_2,\ldots,r_n)=0 \mid p_k(r_2,\ldots,r_n) \neq 0] \leq k/|S|$

We conclude:

$\Pr[\ p(r_1, \ldots, r_n) = 0] \leq (d\text{-}k)/|S| + k/|S| = d/|S|$

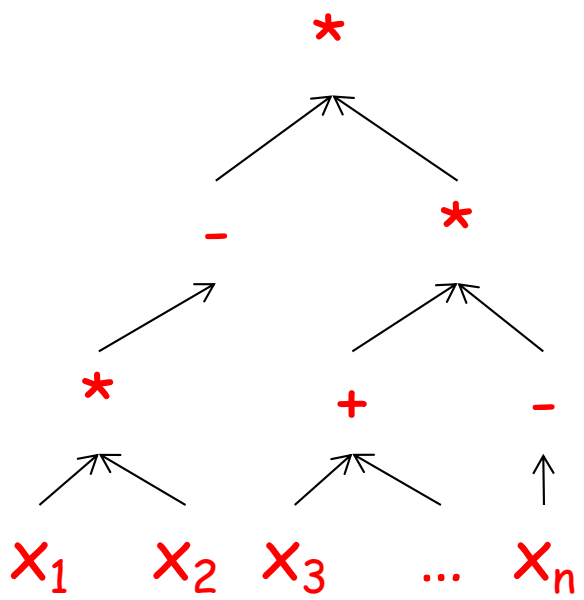
Note: can add these probabilities (1) + (2) because:

$\Pr[E_2] = \Pr[E_2|E_1]\Pr[E_1] + \Pr[E_2|\neg E_1]\Pr[\neg\ E_1]$

$\leq \Pr[E_1] + \Pr[E_2|\neg E_1]$

Schwartz-Zippel Randomized algorithm for testing if a polynomial p is identically zero:

- Given: polynomial $p(x_1, x_2, \ldots, x_n)$
- Is p identically zero?



- Note: degree d is at most the size of input

Schwartz-Zippel Randomized algorithm for testing if a polynomial p is identically zero:

- given field **F**, pick a subset $S \subset \mathbf{F}$ of size 2d
  - pick $r_1, r_2, \ldots, r_n$ from S uniformly at random
  - if $p(r_1, r_2, \ldots, r_n) = 0$, answer "yes"
  - if $p(r_1, r_2, \ldots, r_n) \neq 0$, answer "no"
- if p is identically zero, then never wrong
- if not, Schwartz-Zippel ensures probability of error at most ½

(Can extend the result to multivariate polynomials defined by algebraic circuits.)