Today

- Randomized complexity classes
- Randomized computation
  - Testing polynomial identities.
  - Testing s-t connectivity in undirected graphs.
- Amplification: BPP in $P/_{\text{poly}}$.
- BPP in PH.

Logical terminology

- Completeness: The lowest probability with which instances in L are accepted.
- Soundness (error): The highest probability with which instances not in L are accepted.
- For system to be interesting Completeness must be larger than soundness error. If it is bounded away, have BPP.

Complexity Classes

- ZPP, RP, co-RP, BPP: for zero-sided, one-sided, other-sided, two-sided errors, all in polynomial time.
- ZL, RL, co-RL, BPL: Analogous classes. Catches:
  - Two-input machine has one-way access to random tape.
  - Running time bounded by polynomial (why?).

Testing Polynomial Identities

Will pose as an "oracle" problem:

Given: An oracle $A : \mathbb{Z}^n \to \mathbb{Z}$, such that $A(x_1, \ldots, x_n)$ is a polynomial in n variables of degree $d < \frac{p}{3}$.

Question: Does there exist $x_1, \ldots, x_n$ such that $A(x_1, \ldots, x_n) \neq 0$?

(Warning: Oracle defined for only one input length ... you can extend easily.)

Actually testing if polynomial is zero not if two polynomials are identical; but problems are virtually same.

Algebraic preliminaries

Definitions by example:

Multivariate Polynomials:

$$3x_1^2x_2^3 + x_1^3 - x_2^4$$

is a polynomial in 2 variables x_1 and x_2. Its degree in x_1 is 3, its degree in x_2 is 4 and its total degree is 5 (largest total degree of the monomials in it).

Polynomial identity testing

Relativized problem.

- As posed: in NP^A.
- Will show: in RP^A.
- Exercise: not in P^A.

Many Applications

1. Given Matrix M whose entries are linear functions in $x_1, \ldots, x_n$, determine if the determinant of this matrix is identically zero.
2. Given two "Read-Once-Branching Programs" are they equivalent.

Both problems in RP (or $\mathrm{co\text{-}RP}$), but not known to be in P.

Randomized polynomial identity testing

Algorithm:

- Set $m = 3d$.
- Pick $a_i \in_R \{1, \ldots, m\}$ independently.
- If $A(a_1, \ldots, a_n) \neq 0$ accept, else reject.

Clearly in randomized polynomial time.

Analysis

(Famed Lemma:) If a polynomial p of degree d is non-zero, and S is a finite subset of the domain of the polynomial, then

$$\Pr_{\mathbf{a}\in S^n}[p(\mathbf{a}) = 0] \leq d/|S|.$$

Proof: By Induction.

- Write

  $p(x_1, \ldots, x_n) = x_n^{d_n} q(x_1, \ldots, x_{n-1}) + r(x_1, \ldots,$

  where degree of r in x_n is less than d_n.

- Pick $x_1 = a_1, \ldots, x_{n-1} = a_{n-1}$ first.

- Bad Event E_1: $q(a_1, \ldots, a_{n-1}) = 0$.
- $\Pr[E_1] \leq (d - d_n)/|S|$ (by induction).
- Now assume E_1 does not happen. Let $g(x_n) = p(a_1, \ldots, a_{n-1}, a_n)$. Note degree of g is at most d_n and g is not identically zero.
- Pick $x_n = a_n$ at random now.
- Bad Event E_2: ($\overline{E}_1$ and $g(a_n) = 0$). Note $\Pr[E_2] \leq \Pr[E_2|\overline{E}_1] \leq d_n/|S|$.
- Claim: If E_1 and E_2 don't happen, then $p(\mathbf{a}) \neq 0$.
- Thus $\Pr[p(\mathbf{a}) = 0] \leq \Pr[E_1] + \Pr[E_2] \leq d/|S|$.

USTCON in RL

USTCON: (Undirected S-T CONnectivity):

Given: Undirected graph G and special vertices s and t.

Question: Is there a path connecting s to t?

Clearly USTCON in NL.

Surprisingly in RL.

(Will assume graph is given by adjacency list + vector of degrees.)

Randomized algorithm

1. Initially $u \leftarrow s$. Set time-left $= n^3$.
2. If $u = t$, then halt and accept.
3. If time-left $= 0$ then halt and reject.
4. Else pick random index i in $\{1, \ldots, d_u\}$.
5. Let v to be ith neighbor of u.
6. Let $u \leftarrow v$; decrement time-left; Go to Step 2.

Clearly in RL. Completeness obvious. Soundness?

Blurb on soundness

- Process called a "random walk".
- Special case of "Markov chains": Prob. of future event independent of past history, given current state.
- Random walks are widely studied.
- Mostly well understood. In particular following is known.

Lemma: In undirected connected graph with n vertices, a random walk starting anywhere reaches every vertex in $O(n^3)$ time with probability $2/3$.

(Maybe learn about this is a randomized algorithms course.)

RP Amplification

Suppose M accepts language L with completeness $c(n) = 1/n^2$ (and $s(n) = 0$). How to amplify completeness?

Amplification: Run machine n^4 times on independent random strings $y_1, \ldots, y_{n^4}$, and accept if one of the y_i's accepts.

$$\Pr_{\mathbf{y}}[\exists i \text{ s.t. } M(x, y_i)\text{accepts}] \geq 1 - (1 - 1/n^2)^{n^4} \geq 1\text{-}$$

Thus completeness $1/\text{poly}(n)$ vs. $1 - \exp(n)$ are equivalent.

BPP amplification

- How to use the above idea for BPP?
- Natural idea:
  - Repeat N times.
  - Accept if # acceptances more than $(c + s)N/2$.
- Analysis?
  - Use "tail inequalities".
  - "Chernoff bound".

Chernoff bounds

Suppose $X_1, \ldots, X_N$ are independent identically distributed random variables in the interval $[0,1]$ with $\mathbf{E}[X_i] = \mu$.

Then

$$\Pr[|\frac{1}{N}\sum_i X_i - \mu| \geq \lambda] \leq e^{-\lambda^2 N/2}.$$

Consequence

Let $X_i = 1$ if $M(x, y_i)$ accepts and 0 o.w.

Applying Chernoff bounds, we see that if $N \sim m/(c-s)^2$ then amplification increases completeness to $1 - \exp(-m)$ and decreases soundness to $\exp(-m)$.

Next: Use this to show BPP in $\mathsf{P}/_{\mathrm{poly}}$.

Consequence: BPP in $\mathsf{P}/_{\mathrm{poly}}$

Say $L \in \mathrm{BPP}$. Assume w.l.o.g. that M is a two input machine recognizing L with $c(n) \geq 1 - 4^{-n}$ and $s(n) \leq 1 - 4^{-n}$. (Notice we get this by amplification.)

Say M uses m-bit random strings.

Claim: Exists $r \in \{0,1\}^m$ such that for every x, $M(x, r) = L(x)$.

Proof: Say $y \in \{0,1\}^m$ is BAD for x if $M(x, y) \neq L(x)$.

For any $x \in \{0,1\}^n$ there are at most 2^{m-2n} y's that are BAD for x.

Taking the union of all BAD sets, there are at most 2^{m-n} strings that are BAD for some x.

Since $2^m > 2^{m-n}$ there exists at least one y which is not BAD for any x. Setting $r \leftarrow y$ gives the Claim.

Thm: BPP $\subseteq \mathsf{P}/_{\mathrm{poly}}$.

Proof: $\mathsf{P}/_{\mathrm{poly}}$ machine is M from the argument above. For every n, advice string is the $r \in \{0,1\}^m$ from the claim.

Next: BPP in PH

Note note quite trivial. How to have a bounded round interaction to comvince $x \in L$?

Consider following game: Kasparov & I are all powerful players. I want to convince you (the audience) that $x \in L$ and Gary claims otherwise. How can we prove our claims?

Draw picture here.

Most strings are good (M(x,y) = accept); or very few are good. How to convince you?

Idea 1: I'll divide space into two equal parts with all bad strings in one part and a bijection pi between the two parts. I claim every string

or its map under bijection is good! If Gary wants, he can challenge me!

If Gary finds a string y where neither M(x,y) nor M(x,pi(y)) accept - he wins.

Else I win.

Seems convincing. I can win if bad set is smaller than 1/2. I can't win if bad set more than 1/2.

Problem: How do I give the bijection?

Bijections have to simple: So we'll stick $\pi_r : y \mapsto y \oplus r$.

In this space of bijections the proof doesn't go through. But the idea is starting to emanate.

Debate for membership in BPP

Theorem: If x in L there exist $r_1, \ldots, r_{2m} \in \{0, l\}^m$ such that the y's are covered; i.e., for every y there exists an $i \in [2m]$ such that $M(x, \pi_{r_i}(y))$ accepts.

If x not in L, then for any $r_1, \ldots, r_{2m} \in \{0, l\}^m$ there is an uncovered y.

Assuming theorem: Debate: I announce $r_1, \ldots, r_{2m}$. Gary challenges with a y. You compute $M(x, y \oplus r_1) \vee \cdots \vee M(x, y \oplus r_{2m})$. If true, I win ($x \in L$) else Gary wins ($x \notin L$) - you decide!

Proof of theorem

If x in L

$$\Pr_r[M(x, y \oplus r)] \geq 1 - 2^{-n} \geq 1/2.$$

$$\Pr_{r_1, \cdots, r_{2m}}[\exists i \in [2m] \text{ s.t. } M(x, y \oplus r_i)] \geq 1 - 2^{-2m}.$$

$$\Pr_{r_1, \cdots, r_{2m}}[\forall y \in \{0,1\}^m, \exists i \in [2m] \text{ s.t. } M(x, y \oplus r_i)]$$

Yields first part.

Proof of theorem (second part)

x not in L. Say I pick best possible $r_1, \ldots, r_{2m}$ below.

$$\Pr_y[M(x, y \oplus r_i)] \le 1/100m.$$
$$\Pr_y[\exists i \in [2m] \text{ s.t. } M(x, y \oplus r_i)] \le 1/50.$$

QED!

Power of the prover

If I am right - I just need to pick $r_1, \ldots, r_{2m}$ at random!

If Gary is right, he just needs to pick y at random.

So we just need randomness to simulate randomness!

Hmm.... that didn't sound so impressive - I should have said ...

So we just need one-sided randomness to simulate two-sided randomness! You'll figure out what I mean in problem set!

Current issues in randomness

- Reducing randomness
  - Algorithm specific: Limited independence, Epsilon-bias.
  - Generically, during amplification: "Recycling".
- Using imperfect randomness: Extractors.
- Derandomization: Pseudorandomness, hardness versus randomness.