This week

- Randomized computation.
- Complexity classes.
  - ZPP, RP, co-RP, BPP.
  - ZL, RL, co-RL, BPL.
- Testing polynomial identities.
- Testing s-t connectivity in undirected graphs.
- Amplification: BPP in $P/_{\rm poly}$.
- BPP in PH.

Clarification on Games

Few lectures back .... we said some wrong things.

- Game is in PSPACE only if there is an a priori polynomial upper bound on its running time.
- Go: # of pieces on board increase all the time.
- Geography: Path length bounded by size of Atlas.
- Chess: No "a priori" upper bound - hence not known to be in PSPACE.

Randomized computation

- Physicists' Belief: Natural phenomena have randomness built into them.
- How does this affect our belief that "polynomial time" is all that is feasible?
- Should study formally.

Randomized algorithms/Turing machines

- Model 1: Machine can enter a random state whenever it wishes. Takes one of two outgoing transitions randomly.
- (Equivalent) Model 2: Machine has two inputs: (1) The actual input and (2) the outcome of many independent random coin tosses.

Randomized machines and languages

Machine M for Language L has:

Completeness c if $c = \inf_{x \in L} \Pr_y[M(x,y)\text{accepts}$
(Assume uniform distribution on $\ell(|x|)$ bit strings.

Soundness s if $s = \sup_{x \notin L} \Pr_y[M(x,y)\text{accepts}]$.

M seems to decide membership in L if $c > s$.
But even better if $c = 1$ (and/or $s = 0$).

Complexity Classes

- Resource? Space or Time?
- What kind of error? Two attributes; Four classes.
  - "False positives": Says $x \in L$ while $x \notin L$. (Soundness > 0.)
  - "False negatives": Says $x \notin L$ when $x \in L$. (Completeness < 1.)
- All in all, get eight classes!

Time-bounded randomization

- BPP: (Bounded Probability Polynomial-time): Both kinds of errors allowed (two-sided error): $L \in BPP$ if there exists a two-input deterministic machine M running in time poly in first input such that:

  $$x \in L \Leftrightarrow \Pr_y[M(x,y)\text{accepts}] \geq 2/3.$$

  (Completeness $= 2/3$; Soundness $= 1/3$).
- RP: (Randomized Polynomial-time): Only false negatives (one-sided error):

  $$x \in L \Rightarrow \Pr_y[M(x,y)\text{accepts}] \geq 2/3.$$

  (Completeness $= 2/3$; Soundness $= 0$ (perfect)).

Time-bounded randomization (contd.)

- co-RP: complements of RP languages.
- ZPP: Error happens with probabillity zero! So what does randomness do? Running time is not guaranteed to be polynomial. Only expected to be polytime.

Space-bounded randomization

Similar collection of four classes:

- BPL, RL, co-RL, ZPL.
- Catch 1: In two-input model, have one way access to second input.
- Catch 2: Machines bounded to run in polynomial time.

Looking ahead

- $2/3$, $1/3$ arbitrarily chosen. For definition of BPP suffices to have $c > s$. Similarly for RP, suffices to have $c > 0$ etc.
- Randomness more powerful than deterministic?
  - Belief: No.
  - Current evidence: Yes. There exist problems in RP that we can show to be in P. (Example: Primality testing.) There exist problems in RL that we can't show to be in L. (Example: USTCON - connectivity in undirected graphs.)

Looking further ahead

- How do RP, BPP etc. relate to familiar complexity classes.
- Obviously: ZPP in RP & co-RP; and all are in BPP.
- RP in NP (by definition).
- BPP? Don't quite know:
  - BPP in $P/_{\text{poly}}$.
  - BPP in PH.

Testing Polynomial Identities

Will pose as an "oracle" problem:

Given: An oracle $A : \mathbb{Z}^n \rightarrow \mathbb{Z}$, 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: We're posing problem for only one n, but 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.)

Exercises/Problems

1. Prove bipartiteness in RL.
2. Prove USTCON in ZPL.
3. Prove there is a randomized logspace algorithm that does s-t connectivity in directed graphs (but not running in polynomial time).

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 time: Use this to show BPP in P/$_{\text{poly}}$.