LECTURE 21

Note Title 4/30/2007

Today: Average-Case Complexity: Definitions.

- Distributional Problem ?
- Feasible Problems ?
- Intractable Ones ?
- Reductions ?

Based entirely on:

ODED GOLDREICH: CONCEPTUAL INTRO TO COMPUTATIONAL COMPLEXITY Sect. 10.2

Question:

- Is TSP hard on average or easy?

Answer:

- Depends who you ask!
- If we pick points uniformly from an $n \times n$ square ... then seems easy.
- if you pick entire grid & perturb each point a little, then seems hard.

Conclusion:

Complexity is a function of

(i) Problem &

(ii) Distribution.

Distributional Problems

Specified by a pair (Π, D)

- $\Pi \subseteq \{0,1\}^* \times \{0,1\}^*$: usual relational problem
- $D = \{D_n\}$, $D_n : \{0,1\}^n \to [0,1]$ is a distribution on $\{0,1\}^n$.

—— x ——

Goal: Given $x \leftarrow_D \{0,1\}^n$
find y s.t. $(x, y) \in \Pi$.

—— x ——

Complexity Measure ?

Expected running time ? .... Not so interesting

Examples :

1. Suppose A solves Π on D as follows:
   - w.p. $2^{-\sqrt{n}}$, A takes time 2^n.
   - w.p. $1-2^{-\sqrt{n}}$, A takes time n^2.

   Is this "polynomial"? exponential?

2. Suppose B solves Π' on D' as follows:
   - w.p. $\sim \frac{1}{c^2}$ B takes time n^c.

   Is this "polynomial"?

Our Preference

Avg-Time = "Time as viewed by polytime observer".

& not

"what could be sensed after unreasonable sampling"

Back to Examples:

(i) In any poly # samples, very unlikely to see exponential behavior.

$\Rightarrow$ Avg-Time = n^2.

(ii) for every c, prob. of seeing run time $\geq n^c$, is $> \frac{1}{c^2}$.

$\Rightarrow$ Avg. Time = super-poly

Formal Definition:

Avg-Time of A on (Π, D) is $\leq T(n)$

if $\forall n, c$

$$\Pr_{x \leftarrow_{D_n} \{0,1\}^n} \left[\begin{array}{l} A(x) \text{ incorrect} \\ \text{or } A(x) \text{ runs in time} \\ \quad \geq T(n) \end{array} \right] \leq \frac{1}{n^c}.$$

Note: Allowing A to be incorrect makes definitions equivalent.

$$\text{Avg-BPP} = \left\{ (\Pi, D) \;\middle|\; \begin{array}{l} \exists A, c \text{ solving } (\Pi, D) \\ \text{in Avg-Time } n^c \end{array} \right\}.$$

Intractable Problems?

Attempt 1:

$$DNP_1 = \left\{ (\Pi, D) \;\middle|\; \begin{array}{l} \Pi \in NP, \\ \text{(i.e., "}(x,y) \in \Pi\text{?" decidable in } P\text{)} \\ D \text{ distribution} \end{array} \right\}$$

Problem

- Notion of Distribution too strong, for "empirical" concerns.
- Can easily prove

$$NP \not\subseteq BPP \implies DNP_1 \not\subseteq \text{Avg. } BPP$$

worstcase hardness $\Longrightarrow$ average case-hardness.

- $D^{Adv}_{A,n}$ = uniform on $\{x \mid A(x) \text{ incorrect}\}$

- $D^{Adv}_{n} = \sum_i \frac{1}{i^2} \cdot D^{Adv}_{A_i,n}$

$A_1, A_2, \ldots A_i, \ldots$ enumeration of BPP m/c.

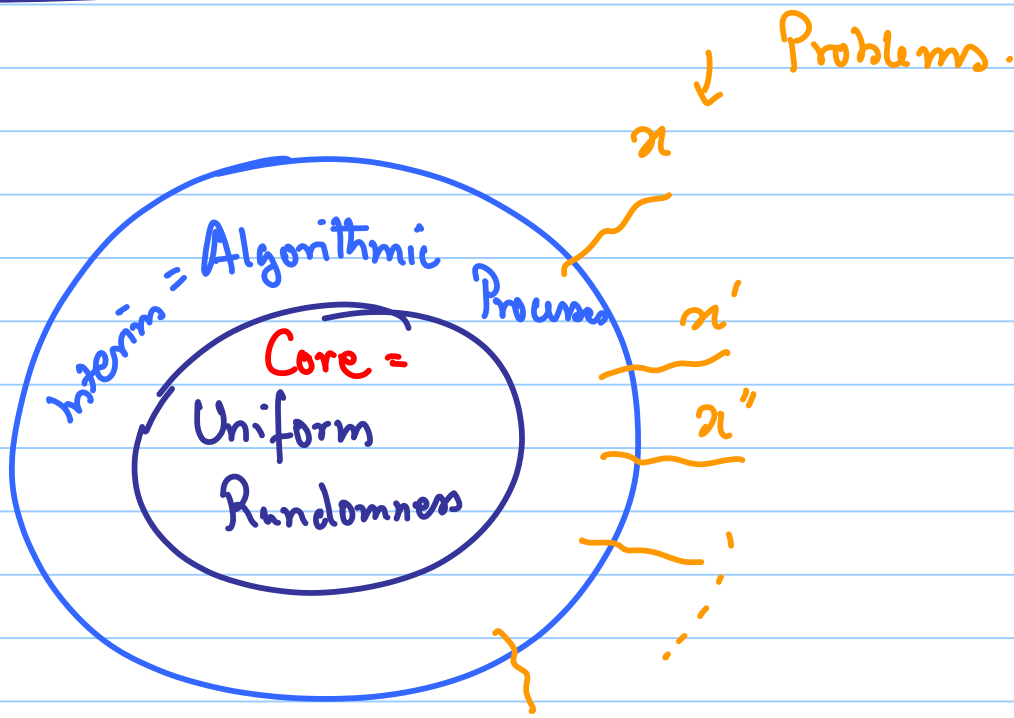
- Diagonalization by Distribution!

- Problem: Distributions worse than adversary, while avg. case wants to understand "naturally-occurring instances".

- Shouldn't allow arbitrary distributions D.

Sampleable Distributions

Model of Universe



- What kind of distribution do we see?
- "Sampleable Distribution"

Definition D is sampleable if

$\exists$ poly time (deterministic) algorithm G

s.t. $\forall$ n, $x \in \{0,1\}^n$

$$\Pr_{y \leftarrow \text{uniform on } \{0,1\}^n} \left[G(y) = x \right] = D(x).$$

Interesting Intractable Problems

- $DNP = \{ (\Pi, D) \mid \Pi \in NP, \ D \text{ sampleable} \}$

Basic Questions

Beliefs

- Is $\mathrm{DNP} \subseteq \text{Avg-BPP}$? No.
- If $\mathrm{NP} \nsubseteq \mathrm{BPP}$ then, Yes, but
  is $\mathrm{DNP} \nsubseteq \text{Avg-BPP}$? can't prove.
- Find some "worst-case assumption" that
  implies $\mathrm{DNP} \nsubseteq \text{Avg. BPP}$.
- What are some DNP-complete
  problems ?
- What is completeness ?
  reductions ?

Reductions

- **Deterministic** : Simple .... should help solve original problem.
- **Probabilistic** : Already get complex ... needn't always be correct.
- **Distributional** : Trickier .... Can be incorrect; can produce unlikely instances.

More formally

Most restrictive notion:

- (Deterministic Reduction): (R, T) reduce

$$(\Pi_1, D_1) \longrightarrow (\Pi_2, D_2)$$

if
(i) R, T are polytime.

(ii) $(R(x), y) \in \Pi_2$

$\Rightarrow (x, T(y)) \in \Pi_1$

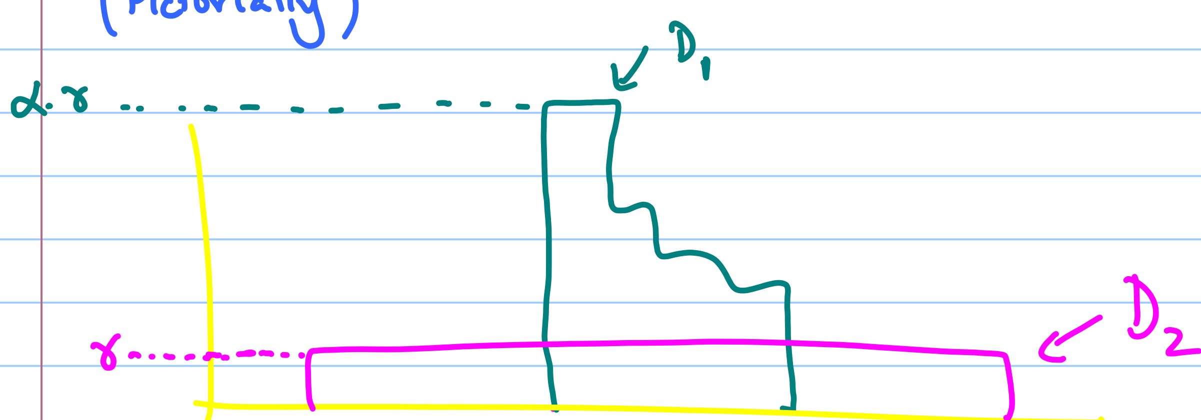
(iii) $R(x)$ distributed as D_2

if x distributed as D_1

- But don't need to adhere to distributions so stringently;
- Alg for Π_2 doesn't "know" D_2.
- <u>Domination of Distributions</u>
- D_2 α-dominates D_1 if

  $\forall x \quad D_1(x) \le \alpha(|x|) \cdot D_2(x)$.

(Pictorially)



- (weaker det. reduction)

$$(\Pi_1, D_1) \longrightarrow (\Pi_2, D_2)$$

$$\begin{tikzcd}
x \arrow[r, "R"] \arrow[d, leftrightarrow, "\Pi_1"] & R(x) \arrow[d, leftrightarrow, "\Pi_2"] \\
T(y) & y \arrow[l, "T"']
\end{tikzcd}$$

$x \leftarrow D_1 \longrightarrow$ $R(x)$ drawn from D_2'

s.t. D_2 poly. dominates D_2'.

__Claim__: Such a reduction + $(\Pi_2, D_2) \in$ Avg. BPP

$\Rightarrow (\Pi_1, D_1) \in$ Avg. BPP.

Can also consider randomized ...

- Definition complex.
- Will see example next lecture.