Structural complexity of **AvgBPP**

Dmitry Itsykson

Steklov Institute of Mathematics at St. Petersburg

CSR 2009
Novosibirsk
August 21, 2009

Outline

(1) Worst-case complexity
  - **BPP**
  - Structural properties: time hierarchy and complete problems

(2) Average-case complexity
  - Distributional problems
  - Average-case tractability

(3) Results: structural properties of **AvgBPP**.

Randomized algorithms with bounded error

- [Gill 1977] Class **BPP** contains languages L decidable by randomized polynomial-time Turing machine M with bounded error: $\forall x \Pr\{M(x) = L(x)\} \geq \frac{3}{4}$
- [Solovay, Strassen 1977] PRIMES$\in$ **BPP**.
- [Agraval, Kayal, Saxena 2002] PRIMES$\in$ **P**.
- Main example: polynomial identity testing.
- [Nisan, Wigderson 1994; ...; Umans 2003] hardness vs randomness tradeoff: $\mathbf{P} = \mathbf{BPP}$ under believable hardness assumption.

Time hierarchy

- A time hierarchy theorem states that a given computational model can decide more languages if it is allowed to use more time.
- [Hartmanis and Stearns 1965] $\mathbf{DTime}[n^a] \subsetneq \mathbf{DTime}[n^{a+\epsilon}]$.
- [Cook 1972] $\mathbf{NTime}[n^a] \subsetneq \mathbf{NTime}[n^{a+\epsilon}]$.
- [Karpinski and Verbeek 1987] $\mathbf{BPTime}[n^{\log n}] \subsetneq \mathbf{BPTime}[2^{n^\epsilon}]$.
- [Barak 2002; Fortnow, Santhanam 2004; van Melkebeek, Pervyshev 2007] Time hierarchy for $\mathbf{BPP}/\mathbf{1}$.
- [Fortnow, Santhanam 2004; Pervyshev 2007] Time hierarchy for heuristic **BPP**.
- We can't prove that $\mathbf{BPTime}[n] \neq \mathbf{BPP}$

Complete problems

B is a complete problem for the class **C** if $B \in \mathbf{C}$ and $\forall A \in \mathbf{C}$, A reduces to B.

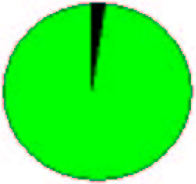
- [Cook, Levin, 1971] **NP**-complete problems: Bounded Halting, Tiling, SAT, TSP,...
- Complete problem for **BPP** is not known.
- **BPP**-complete language $\implies$ time hierarchy for **BPP**.
- [Hartmanis, Hemachandra 1986] $\exists$ oracle A, such that $\mathbf{BPP}^A$ doesn't have complete languages.
- $\mathbf{P} = \mathbf{BPP} \implies$ **BPP** has complete language.
- Time hierarchies and complete problems usually require enumeration of (correct) machines in the respective computational model. We don't know how to enumerate machines that have bounded error.

Average-case tractability

- Distribution $D = \{D_n\}_{n=1}^{\infty}$ where $D_n : \{0,1\}^n \to \mathbb{R}_+$ such that $\sum_{a \in \{0,1\}^n} D_n(a) = 1$.
- Distributional problem (L, D), where L is a language, D is a distribution.
- Polynomial-time samplable distribution $\exists$ polynomial time randimized algorithm (sampler) S such that $S(1^n)$ is distributed according D_n.

Levin (1986):
$T(x)$ is runing time on input x;
$T(x)$ is polynomial on the average if
$\exists \epsilon > 0 : \mathrm{E}_{x \leftarrow D_n} T^{\epsilon}(x) = O(n)$

Typical situation:
- $\frac{1}{\exp}$: exponential time
- $1 - \frac{1}{\exp}$: polynomial time



AvgP, AvgBPP

Class	Problem	Turing machine	Time	Error
P	language L	deterministic M	poly	no error $\forall x\ M(x) = L(x)$
BPP	language L	randomized M	poly	bounded error $\forall x \Pr[M(x) = L(x)] \geq \frac{3}{4}$
AvgP	distr. problem (L, D)	deterministic M	avg. poly	no error $\forall x\ M(x) = L(x)$
Avg-BPP	distr. problem (L, D)	randomized M	avg. poly	bounded error $\forall x \Pr[M(x) = L(x)] \geq \frac{3}{4}$

Results

1. Proper inclusions:
   - $\mathbf{P} \subsetneq \mathbf{AvgP} \subsetneq \mathbf{EXP}$;
   - $\mathbf{BPP} \subsetneq \mathbf{AvgBPP} \subsetneq \mathbf{BPEXP}$.
2. Time hierarchy theorem for $(\mathbf{AvgBPP}, \mathbf{PSamp})$.
3. Construction of distributional problem (C, R) that is complete in $(\mathbf{AvgBPP}, \mathbf{PSamp})$ under deterministic Turing reduction.
   - If $(C, R) \in \mathbf{AvgP}$, then $(\mathbf{AvgP}, \mathbf{PSamp}) = (\mathbf{AvgBPP}, \mathbf{PSamp})$
   - R is enough complicated samplable distribution.
   - Existence of complete problem with uniform (or uniform-like) distribution implies some derandomization $(\mathbf{BPEXP} \subseteq \mathbf{AvgEXP})$.

Why do we fail with **BPP**-complete problem?

- $X = \{(M, x, 1^t) \mid M$ is a bounded error randomized TM, $\Pr[M^{\leq t}(x) = 1] \geq \frac{3}{4}\}$
- Let L be solvable in **BPP** by TM M in n^c steps.
- $x \in L \iff (M, x, 1^{n^c}) \in X$.
- X is **BPP** hard, X is probably not decidable.
- $Y = \{(M, x, 1^t) \mid M$ is a randomized TM, $\Pr[M^{\leq t}(x) = 1] > \frac{1}{2}\}$
- Y is **BPP**-hard, decidable, $Y \stackrel{?}{\in}$ **BPP**.

Idea of **AvgBPP** complete problem

- $Y = \{(M, x, 1^m) \mid M \text{ is a rand. TM}, \Pr[M^{\le t}(x) = 1] > \frac{1}{2}\}$
- freq $M^{\le t}(x)$ is the most frequent answer of $M^{\le t}(x)$.
  prob $M^{\le t}(x) = \Pr[M^{\le t}(x) = \text{freq } M^{\le t}(x)]$

Good	Intermediate	Bad
$(M, x, 1^t)$: prob $M^{\le t}(x) \ge \frac{3}{4}$	$(M, x, 1^t)$: $\frac{3}{4} >$ prob $M^{\le t}(x) \ge \frac{3}{5}$	$(M, x, 1^t)$: prob $M^{\le t}(x) < \frac{3}{5}$

- If $(M, x, 1^t)$ is good or intermediate, then there is poly-time randomized test that checks $\Pr[M^{\le t}(x) = 1] > \frac{1}{2}$.
- Just repeat executions many times and output the most frequent answer.
- If prob $M^{\le t}(x) \approx \frac{1}{2}$, then we have to go through all sequences of random bits.

Distributional problem

- Idea: to construct such samplable distribution R that bad $(M, x, 1^t)$ will have R-mesure $2^{-\Omega(n^2)}$.

Sampler $\mathcal{R}(1^n)$

1. Generate: (M, x, t), $|(M, x, 1^t)| = n$.
2. Run Test
   - Run $M^{\leq t}(x)$ for n^2 times.
   - If at least 0.7 fraction of executions output the same value, then return (M, x, t).
   - Return 0^n.

- R-measure of all bad $(M, x, 1^t)$ is bounded by $2^{-\Omega(n^2)}$.
- R-measure of good $(M, x, 1^t)$ is $\approx 2^{-|M|-|x|-\log t}$.

$(Y, R) \in$ **AvgBPP**

1. Run $M^{\leq t}(x)$ for n^2 times.
2. If at least 0.65 fraction of executions output the same value, then return it's value.
3. Otherwise w.h.p. R-measure of $(M, x, 1^t)$ is $2^{-\Omega(n^2)}$. We just go through all sequences of random bits and compute the answer deterministically.

- Y is **BPP**-hard.
- $(Y, R) \in$ **AvgBPP**.

Open questions

- Time hierarchy theorem and complete problem for **AvgRP**.
- To classify properties that we can prove in average case and can't prove in worst case.