

Lecture 19: $BPP \subseteq \Sigma_2^P \cap \Pi_2^P$

Instructor: Jin-Yi Cai

Scribe: Rakesh Kumar, Yunpeng Li

Last time we proved by the Adleman's argument that $BPP \subseteq P/poly$ which relates BPP to the nonuniform complexity class. Today we will study the relationship between BPP and the uniform complexity class.

THEOREM 0.1 (SIPSEK-GAĆS, LAUTEMANN).

$$BPP \subseteq \Sigma_2^P \cap \Pi_2^P$$

The result that BPP is contained in the polynomial-time hierarchy was given by Michael Sipser, where Sipser originally announced that $BPP \subseteq \Sigma_4^P \cap \Pi_4^P$, then he improved it to $BPP \subseteq \Sigma_3^P \cap \Pi_3^P$. Gaćs noticed that Sipser's proof actually could get $BPP \subseteq \Sigma_2^P \cap \Pi_2^P$. A simpler version of this theorem was published later by Lautemann. We will present both proof. In addition, there is another proof which explores the similarity between probabilistic and nondeterministic Turing machine and to establish a polynomial qualifier characterization of the class BPP and get a *Swapping Lemma*. (See the textbook for more formal treatment.)

1 Lautemann's Proof

In this section, we present Lautemann's proof by first introducing a technical lemma. We examine some properties based on the density of a set.

DEFINITION 1.1 (FAT/THIN SET). Let $S \subseteq \{0, 1\}^m$,

$$S \text{ is fat if } |S| \geq (1 - \frac{1}{m})2^m$$

and we call S is thin if its complement set S^c is fat.

Note that the parameter here could be arbitrarily set according to the requirement. The gist of this definition is that there is a significant *gap* between the cardinality. Suppose we have a set $S \subseteq \{0, 1\}^m$, we regard $\{0, 1\}^m$ as the m -dimensional vector space. Pick any $z \in \{0, 1\}^m$ and define the shift of S by

$$S + z = \{v \oplus z | v \in S\}$$

where \oplus denotes the addition over \mathbb{Z}_2^m . It is obvious that $|S| = |S + z|$.

We want to know whether a small number of shifts to S would cover the whole thing, $\{0, 1\}^m$. The following lemma tells us that it is very probable if S is fat and it is impossible when S is thin.

LEMMA 1.2. Let $S \subseteq \{0, 1\}^m$ then

- If S is thin, for every $r \leq m$

$$Pr_{\mu_1, \mu_2, \dots, \mu_r} [\bigcup_{i=1}^r (S + \mu_i) = \{0, 1\}^m] = 0$$

- If S is fat, for every $r \geq m$

$$Pr_{\mu_1, \mu_2, \dots, \mu_r} \left[\bigcup_{i=1}^r (S + \mu_i) = \{0, 1\}^m \right] \geq 1 - \frac{2^m}{m^m}$$

where the condition $\bigcup_{i=1}^r (S + \mu_i) = \{0, 1\}^m$ means that $(\forall v \in \{0, 1\}^m)(\exists i \leq r)[v \in S + \mu_i]$.

Proof. • We note that $|S + \mu_i| = |S| < 2^m/m$. So

$$\left| \bigcup_{i=1}^r (S + \mu_i) \right| < 2^m$$

Thus,

$$\bigcup_{i=1}^r (S + \mu_i) \neq \{0, 1\}^m.$$

- When S is a fat set, consider the probability

$$\begin{aligned} & Pr_{\mu_1, \dots, \mu_r} \left[\bigcup_{i=1}^r (S + \mu_i) \neq \{0, 1\}^m \right] \\ &= Pr_{\mu_1, \dots, \mu_r} [\exists y \in \{0, 1\}^m \text{ and } y \notin \bigcup_{i=1}^r (S + \mu_i)] \\ &\leq \sum_{y \in \{0, 1\}^m} Pr_{\mu_1, \dots, \mu_r} [y \notin \bigcup_{i=1}^r (S + \mu_i)] \\ &\leq \sum_{y \in \{0, 1\}^m} Pr_{\mu_1, \dots, \mu_r} \left[\bigwedge_{i=1}^r (y \notin S + \mu_i) \right] \\ &\leq \sum_{y \in \{0, 1\}^m} \prod_{i=1}^r Pr_{\mu_i} [y \notin S + \mu_i] \leq \frac{2^m}{m^r} \end{aligned}$$

The last two inequalities derive as follows. Fix a y , the event that $y \notin S + \mu_i$ is independent for every randomly chosen μ_i and $y \notin S + \mu_i$ is equivalent to $\mu_i \notin S + y$. Since y is fixed $|S + y| \geq (1 - \frac{1}{m})2^m$, the probability of y is not in $S + \mu_i$ is at most $1/m$. Since $r \geq m$, the lemma follows. □

We consider the dual version of this lemma by focusing on whether a small number of shifts have nonempty intersection:

LEMMA 1.3. *Let $S \subseteq \{0, 1\}^m$ then*

- If S is thin, for every $r \geq m$

$$Pr_{\mu_1, \mu_2, \dots, \mu_r} \left[\bigcap_{i=1}^r (S + \mu_i) \neq \emptyset \right] \leq \frac{2^m}{m^r}$$

- If S is fat, for every $r \leq m$,

$$Pr_{\mu_1, \mu_2, \dots, \mu_r} \left[\bigcap_{i=1}^r (S + \mu_i) \neq \emptyset \right] = 1.$$

Now let's consider the BPP languages. Let $L \in \text{BPP}$, and the probabilistic Turing machine M

$$\begin{aligned} x \in L &\Rightarrow Pr_{y'}[M(x, y') = 1] \geq \frac{2}{3} \\ x \notin L &\Rightarrow Pr_{y'}[M(x, y') = 1] \leq \frac{1}{3} \end{aligned}$$

We can amplify it such that

$$x \in L \Rightarrow Pr_{y \in \{0,1\}^m} [M(x, y) = 1] \geq 1 - \frac{1}{m} \quad (1)$$

$$x \notin L \Rightarrow Pr_{y \in \{0,1\}^m} [M(x, y) = 1] \leq \frac{1}{m} \quad (2)$$

The witness of the BPP language L is defined as

DEFINITION 1.4 (WITNESS SET). *Let $L \in \text{BPP}$ and M is the probabilistic Turing machine satisfies Equation (1) and (2), the witness set of x is defined as*

$$W_x = \{y | M(x, y) = 1\}$$

Note that the witness set according to our amplification version of BPP is either fat ($x \in L$) or thin ($x \notin L$). Combined with the above two technical lemma, we get:

- if $x \in L$, since the witness set W_x is fat

$$(\exists \mu_1, \dots, \mu_m)(\forall y \in \{0, 1\}^m)[\exists 1 \leq i \leq m, y \in W_x + \mu_i]$$

- if $x \notin L$, since the witness set W_x is thin

$$(\forall \mu_1, \dots, \mu_m)(\exists y \in \{0, 1\}^m)[\forall 1 \leq i \leq m, y \notin W_x + \mu_i]$$

The above two conditions are complement of each other, The predication $\exists 1 \leq i \leq m y \in W_x + \mu_i$ can be checked in polynomial time by letting

$$P(x, \mu_1, \dots, \mu_m, y) = \bigvee_{i=1}^m M(x, y \oplus \mu_i)$$

So we have a Σ_2^P machine for L . Since BPP is closed under complement, then $\text{BPP} \subseteq \Sigma_2^P \cap \Pi_2^P$.

2 Sipser-Gaács' Proof

We introduce a useful concept called isolation.

DEFINITION 2.1 (ISOLATION). *Let \mathcal{H} be a set of 2-universal hash functions from \mathcal{U} to T and S be a subset of \mathcal{H} . We say that a hash function $h \in \mathcal{H}$ isolates an element $z \in S$ if*

$$(\forall z' \in S, z \neq z')[h(z) \neq h(z')].$$

We call a set $\Gamma = \{h_1, \dots, h_r\}$, with $h_i \in \mathcal{H}$, isolates z if there exists a $h \in \Gamma$ that isolates z .

The intuition of the following lemma is that if S is bigger than T , Γ cannot isolate all the elements in S while if S is much smaller than T , with high probability, Γ could isolate each element in S , when each h_i is picked randomly from \mathcal{H} .

LEMMA 2.2 (ISOLATION LEMMA). *Let \mathcal{H} be a family of 2-universal hash functions from \mathcal{U} to T with $\mathcal{U} = \{0, 1\}^m$ and $S \subseteq \mathcal{U}$. Pick $h_1, \dots, h_r \in \mathcal{H}$ uniformly and independently, and let $\Gamma = \{h_1, \dots, h_r\}$*

- if $\frac{|S|}{|T|} \geq r$,

$$Pr_{h_1, \dots, h_r}[\Gamma \text{ isolates all elements in } S] = 0$$

- if $|S|^{r+1} \leq |T|^r$,

$$Pr_{h_1, \dots, h_r}[\Gamma \text{ isolates all elements in } S] \geq 1 - \frac{|S|^{r+1}}{|T|^r}$$

We will prove this lemma next time. We now sketch a proof of Theorem 0.1. Consider a BPP machine M . We will assume that the machine has been suitably amplified. A set T of suitable size will be chosen. For an input x , we will apply the lemma to the witness set W_x . When $x \in L$, $|W_x|$ will be large enough compared to $|T|$ so that no set of r of hash functions can isolate all the elements in W_x . On the other hand, if $x \notin L$, $|W_x|$ will be small enough compared to $|T|$ so that a set of r randomly chosen hash functions will isolate all elements in W_x . Here r is some parameter that will be fixed appropriately so that the above claim is true. Then, we get $x \notin L$ if and only if there exists a set of r hash functions that isolate all elements in W_x . This predicate can be written as

$$(\exists h_1, \dots, h_r)(\forall z \in W_x)(\exists 1 \leq i \leq r)(\forall z' \in W_x, z \neq z')[h_i(z) \neq h_i(z')]$$

Note that the third qualifier $\exists 1 \leq i \leq r$ is of polynomial size and can be eliminated and then we can collapse the consecutive \forall qualifiers to obtain a Π_2^P machine for L .