CS 229r Information Theory in Computer Science Mar 31, 2016

Lecture 18

Lecturer: Madhu Sudan *Scribe: Victor Balcer*

Today will be our first day on the parallel repition theorem. In this lecture, we will set up the background with 2-prover 1-round games, state the theorem and a key lemma that we will use to prove the theorem in the following class.

1 2-Prover 1-Round Games

We will start with an example. Given a nonbipartite graph $G = (V, E)$ (for simplicity we consider an odd cycle). There are two provers Alice and Bob (A and B) who try to convice a third player (which we will call the verifier) that the graph G is 2-colorable. We consider the graph G to be public, but large enough that the verifier is unwilling to check for itself whether the graph is 2-colorable. Because it is nonbipartite, G is not 2-colorable. We set up the game in the following way:

- The verifier picks $x, y \in V$ and gives x to Alice and y to Bob (Alice does not know Bob's input and vice versa).
- Alice and Bob will return to the verifier a color based on some 2-coloring. We denote the outputs $a = A(x)$ and $b = B(y)$. Based on these responses the verifier will decide whether or not the graph is 2-colorable.
- To catch Alice and Bob in a lie the verifier must choose the verticies in a strategic manner. It can either
  1. pick $x = y \in V$ where it would expect Alice and Bob to return the same color or
  2. pick $(x, y) \in E$ where it would expect the returned colors to be different.

So the best strategy for Alice and Bob to cheat the verifier would be to agree on some maximal 2-coloring such that exactly one edge has the same color verticies and the other edges are alternating in color. We say the verifier will choose values randomly (the case to test and then the actual verticies). The verifier will catch the lie only if it queries the one edge with the same color verticies. Therefore, Alice and Bob win the game (convince the verifier the graph is 2-colorable) with probability $1 - \frac{1}{2|E|}$ ($\frac{1}{2}$ for the edge case and $\frac{1}{|E|}$ for the bad edge).

We now give the general definition for 2-prover 1-round games.

Definition 1 *We define a* **2-prover 1-round game** *(call it G) as follows. There are two provers A and B which are given as inputs x and y respectively. x and y are sampled from distributions X and Y (not necessarily independent). The goal of the provers is to satisfy a predicate $V(x, y, a, b) = 1$.*

$$a \leftarrow A \leftarrow x \leftarrow X \qquad\qquad b \leftarrow B \leftarrow y \leftarrow Y$$

In the context of our example

$$V(x, y, a, b) = \begin{cases} 1 & (x = y) \wedge (a = b) \\ 1 & (x \neq y) \wedge (a \neq b) \\ 0 & \text{otherwise} \end{cases}$$

We are interested in the how well the provers can achieve their goal.

Definition 2 *We define the* **game value** *of G as*

$$\mathrm{val}(G) = \max_{A,B} \Pr_{\substack{x \leftarrow X \\ y \leftarrow Y}} [V(x, y, a, b) = 1]$$

where the maximum is taken over strategies A and B.

1.1 Repitition

We can increase the probability of catching a lie by repeating the game multiple times. If in all instances Alice and Bob wins, then the verifier should be convinced Alice and Bob are telling the truth. If they lose any instance, then the verifier has caught them in a lie. We can modify our game to capture repitition with two models.

Consider repitition by repeating the game multiple times sequentially; provers are given the next input after responding to the current input. This is **serial repitition** and has game value $\mathrm{val}(G)^t$ where t is the number of rounds.

Consider a different model where the inputs and outputs of Alice and Bob are now replaced with tuples.

$$(a_1, \ldots, a_t) \leftarrow A \leftarrow (x_1, \ldots, x_t) \leftarrow X \qquad\qquad (b_1, \ldots, b_t) \leftarrow B \leftarrow (y_1, \ldots, y_t) \leftarrow Y$$

In this setting the provers get all inputs at the same time. So the provers win if and only if they win every round. This model is known as **parallel repitition** and will be written as G^t.

Definition 3 *The* **game value** *of G^t is defined as*

$$\mathrm{val}(G^t) = \max_{A,B} \Pr[(\forall i \text{ s.t. } 1 \leq i \leq t)(V(x_i, y_i, a_i, b_i) = 1)]$$

2 Parallel Repitition Theorem

It was originally conjectured that $\mathrm{val}(G^t) = \mathrm{val}(G)^t$. However this is incorrect. Trivially $\mathrm{val}(G^t) \geq \mathrm{val}(G)^t$ because we can consider each input independent from the other inputs (t repititions of G). The following counterexample shows Alice and Bob can do better than $\mathrm{val}(G)^t$.

We construct the Feige game [1] as follows:

1. $x, y \leftarrow \{0, 1\}$ (independent and uniformly drawn) are given to Alice and Bob, respectively.
2. Alice and Bob have outputs of the form $a = (P_a, \beta_a)$ and $b = (P_b, \beta_b)$ where $P \in \{A, B\}$ and $\beta \in \{0, 1\}$.
3. The verifier has predicitate $V(x, y, a, b) = 1$ if and only if $a = b = (A, x)$ or $a = b = (B, y)$.

Intuitively, Alice and Bob win if and only if they agree on the same player (A or B) and they both correctly predict the input given to that player.

Claim 4 *The Feige game G has game values*

$$\mathrm{val}(G) = \frac{1}{2} \qquad \textit{and} \qquad \mathrm{val}(G^2) = \frac{1}{2}$$

Proof $\text{val}(G) = \frac{1}{2}$ because no matter the strategy one player must always guess the other player's input.

Consider the game G^2. Alice has input x_1, x_2 and Bob has input y_1, y_2. The players will output

$$\begin{aligned} a_1 &= (A, x_1) \quad b_1 = (A, y_2) \\ a_2 &= (B, x_1) \quad b_2 = (B, y_2) \end{aligned}$$

By construction, Alice and Bob will win if and only if $x_1 = y_2$. Let w_i be the event that Alice and Bob win the i-th game. So

$$\Pr[w_1 w_2] = \Pr[w_1]\Pr[w_2|w_1] = \frac{1}{2} \cdot 1 = \frac{1}{2}$$

The first probability is 1/2 because the inputs are drawn randomly and the second probability is 1 as conditioning on winning the first game ensures $x_1 = y_2$. ■

So this counterexample has shown us that $\text{val}(G^t) \not\leq \text{val}(G)^t$. Now, we want to know what is a good upperbound for $\text{val}(G^t)$?

Claim 5 $\text{val}(G)$ *is a naive upperbound to* $\text{val}(G^t)$*.*

Proof Assume otherwise. Then there exists some strategy where $\Pr[w_1, \ldots, w_t] > \text{val}(G)$. So

$$\Pr[w_1] \cdot \ldots \cdot \Pr[w_t | w_1, \ldots, w_{t-1}] > \text{val}(G)$$

This implies $\Pr[w_1] > \text{val}(G)$, a contradiction. ■

Theorem 6 (Verbitsky [2]) *If* $\text{val}(G) < 1$*, then*

$$\lim_{t\to\infty} \text{val}(G^t) = 0$$

However this result is too weak for application as the decay is too slow as it depends on the size of the game (the number of possible question answer pairs). In addition, the proof is existential and does not yield an actual strategy.

In the context of our graph coloring example, this theorem implies that the verifier is likely to catch the lie if the game is repeated many times (exponential in the number of verticies), but not for a fixed number of times independent of the graph size.

2.1 Main Theorem

The parallel repitition theorem is stated below. It was originally proved by Raz in 1998 and simplified by Holenstein in 2007 [3, 4].

Theorem 7 *If* $\text{val}(G) < 1 - \alpha$*, then*

$$\text{val}(G^t) \leq (1 - \alpha^3)^{\Omega_{k,\alpha}(t)}$$

where k is the number of responses.

We will use the notation w_i for the event where $V(x_i, y_i, a_i, b_i) = 1$ and w_S for the event where $V(x_i, y_i, a_i, b_i) = 1$ for all $i \in S \subset \{1, \ldots, t\}$. The following lemma will be an important part of our proof.

Lemma 8 (Raz's Lemma) *If* $\mathrm{val}(G) < 1 - \alpha$ *then there exists a constant* $\gamma(k, \alpha)$ *such that* $\forall S \subset \{1, \ldots, t\}$ *with* $|S| \leq \gamma t$*, then either*

1. $\Pr[w_S] \leq 2^{-\gamma t}$

2. $\exists i \notin S$ *such that* $\Pr[w_i | w_S] \leq 1 - \alpha/2$

Assuming this lemma, the previous theorem follows.

Proof [Theorem 7] We start with an empty set S_0 and will continue adding verticies as long as we can keep applying the previous lemma. We can win S_0 with probability 1 as it is empty. Then there exists i_1 such that $\Pr[w_{i_1} | w_{S_0}] \leq 1 - \alpha/2$. Thefefore, $S_1 = S_0 \cup \{i_1\}$. We can keep repeating this process

1. until we have a set S_j such that $\Pr[w_{S_j}] \leq 2^{-\gamma t}$. Thus we are done as winning the entire game requires us to win on S_j, but that happens with exponentially small probability.

2. or $|S_j| = \gamma t$. But from this step, we can compute the probability of winning.

$$\Pr[w_{S_j}] = \Pr[w_{i_j} | w_{S_{j-1}}] \Pr[w_{S_{j-1}}] \leq (1 - \alpha/2) \Pr[w_{S_{j-1}}] \leq (1 - \alpha/2)^{\gamma t}$$

Therefore, $\mathrm{val}(G^t) \leq \max\{2^{-\gamma t}, (1 - \alpha/2)^{\gamma t}\}$ which completes the proof. ■

2.2 Proof Idea for Raz's Lemma

For now, we will provide a high level sketch for the proof of Raz's Lemma. We will prove the lemma by showing the contrapositive. i.e. if there exists S such that $|S| \leq \gamma t$, $\Pr[w_S] \geq 2^{-\gamma t}$ and $\forall i \notin S$ $\Pr[w_i | w_S] \geq 1 - \alpha/2$, then $\mathrm{val}(G) \geq 1 - \alpha$.

We proceed by considering some instance of G. The provers will then embed G into an instance of G^t in a way that allows the provers to win on the embedded instance with high probability which completes the proof. This strategy is similar to what we covered in an earlier lecture on set disjointness.

2.2.1 Randomized Provers

Up until now we have been using deterministics provers. In the proof we will consider randomized provers even though both of these are models are equivalent. Let r be public randomness, r_A and r_B be the private randomness for Alice and Bob. Therefore, A_{r, r_A} is deterministic when the randomness is fixed (likewise for Bob). Therefore,

$$\mathrm{val}(G) = \mathop{\mathbb{E}}_{r, r_A, r_B} \left[\mathop{\mathrm{val}}_{A_{r, r_A}, B_{r, r_B}} (G) \right]$$

So we can just fix the provers that maximize this value which completes the argument for equivalence.

2.2.2 Main Idea

We start with an instance of G (inputs x and y to Alice and Bob). Because we plan to embed in a higher dimension Alice and Bob need to be strategies on inputs $(x_1, \ldots, x_t)$ and $(y_1, \ldots y_t)$. For any fixed strategy we assume that Raz's Lemma is false. i.e. there exists S such that $|S| \leq \gamma t$, $\Pr[w_S] \geq 2^{-\gamma t}$ and $\forall i \notin S \Pr[w_i | w_S] \geq 1 - \alpha/2$.

To understand conditioning on w_S we consider all inputs where Alice and Bob win on S (their strategies need to be fixed to do this). Note that winning on S may restrict the inputs on coordinates not in S. In the case of the Feige game, $y_2 \not\sim (y_2 | x_1, w_1)$ because winning on the first game requires $x_1 = y_2$.

We will need to use the following lemma.

Lemma 9 (Sublemma) *Given the distribution*

$$\left(\begin{array}{c} x_1, \ldots, x_t \\ y_1, \ldots, y_t \end{array} \middle| w_S \right)$$

there exists $i \notin S$ such that the distributions $(x_i, y_i | w_S) \sim (x_i, y_i)$.

In the context of the Feige example the distributions $(x_2, y_2 | w_1)$ and (x_2, y_2) are identical as w_1 only gives the condition $x_1 = y_2$.

So imagine there is some god looking at (x, y) who will create $(x_1, \ldots, x_t)$ and $(y_1, \ldots y_t)$ for Alice and Bob. The god will sample from the distribution

$$\left(\begin{array}{c} x_1, \ldots, x_t \\ y_1, \ldots, y_t \end{array} \middle| w_S \right)$$

while setting (x, y) as the inputs to the i-th game (i from the sublemma). The god can set the i-th game to the original inputs because the sublemma states that conditioning on w_S does not change the distribution of the inputs to the i-th game. By our assumption that Raz's Lemma is false and $i \notin S$ we have $\Pr[w_i | w_S] \geq 1 - \alpha/2$. Now, we assume that Alice and Bob know i and S.

So given this god Alice and Bob have a strategy to win the i-th game (original inputs) with probability $\geq 1 - \alpha/2$. However, Alice and Bob cannot sample from this god distribution exactly and the proof will show that the error of simulating god is small ($< \alpha/2$). Thus, we can still win the original game with probability $\geq 1 - \alpha$ which completes the proof.

2.2.3 Difficultly with Simulating God

Alice and Bob are able to use shared randomness to sample from the god distribution except on the i-th game because they do not know each other's inputs. Informally, to simulate god Alice and Bob both want to sample from the distribution

$$\left(\begin{array}{c} x_1, \ldots, x_t \\ y_1, \ldots, y_t \end{array} \middle| w_S, \begin{array}{c} x_i = x \\ y_i = y \end{array} \right)$$

However, they only know their own input and must sample from (Alice on the left and Bob on the right)

$$\left(\begin{array}{c} x_1, \ldots, x_t \\ y_1, \ldots, y_t \end{array} \middle| w_S, x_i = x \right) \qquad \left(\begin{array}{c} x_1, \ldots, x_t \\ y_1, \ldots, y_t \end{array} \middle| w_S, y_i = y \right)$$

Part of the proof will show all three of these distributions are "close" and therefore we can use a correlated sampling technique to simulate sampling from our god distribution with minimal error.

References

[1] Feige, Uriel. On the success probability of the two provers in one-round proof systems. In *Structure in Complexity Theory Conference, Proceedings of the Sixth Annual.* IEEE, 1991.

[2] Verbitsky, Oleg. Towards the parallel repitition conjecture. In *Structure in Complexity Theory Conference, Proceedings of the Sixth Annual.* IEEE, 1994.

[3] Raz, Ran. A parallel repetition theorem. *SIAM Journal on Computing 27.3*, 1998.

[4] Holenstein, Thomas. Parallel repetition: simplifications and the no-signaling case. *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing.* ACM, 2007.