CS 229r Information Theory in Computer Science Apr 4, 2019

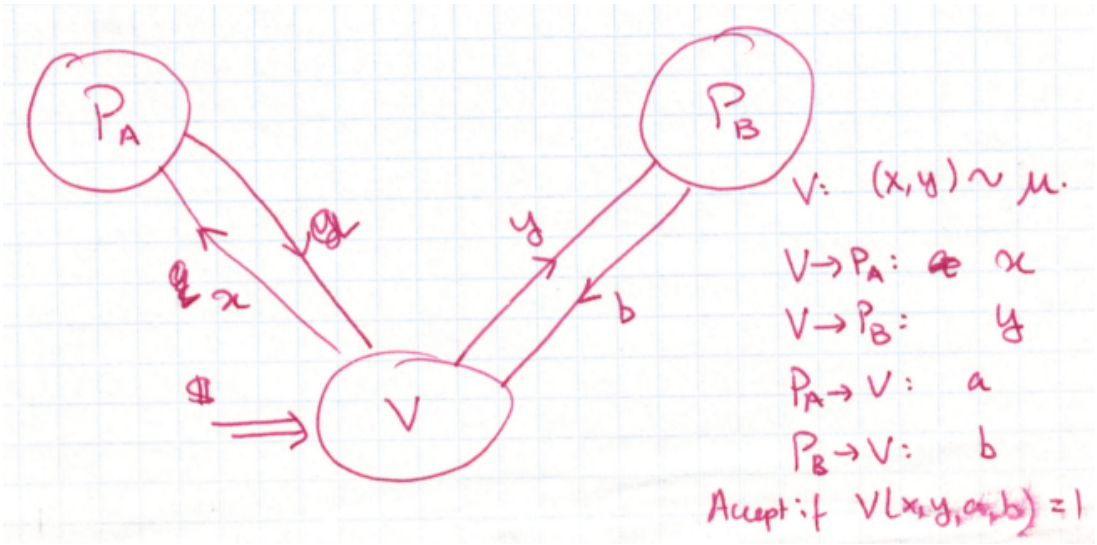
Lecture 1

Instructor: Madhu Sudan *Scribe: Matthew Hase-Liu*

Today, we will begin studying the parallel repetition theorem, beginning with 2-prover games, followed by motivation and examples, the repetition problem and question, and finally the main theorem.

2-Prover Games

We have three players, two of which are provers and one of which is a verifier. In the diagram below, we denote the provers Alice and Bob, say A and B, respectively.



In particular, there exactly one round of interaction. We have the following procedure:

(i) The verifier first selects (x, y) according to some distribution μ.

(ii) Alice receives x and Bob receives y, and neither prover knows the other's input.

(iii) Alice and Bob then each return to the verifier a and b, respectively.

(iv) The verifier chooses to accept or reject, depending on x, y, a, b.

The simplest example of this is the "odd cycle game":

Example 1 (Odd Cycle Game)**.** As a rough idea, the provers are claiming that $C_n = (\mathbb{Z}_n, E = \{i \bmod n, i+1 \bmod n\})$ is 2-colorable (n odd). Here is a possible protocol:

(i) Select (x, y) distributed as follows:
With probability $1/2n$, select $(i, i+1)$ (over all possible n values of i), and with probability $1/2n$, select (i, i) (also over all possible n values of i).

(ii) Alice then sends back to the verifier some $X_A(x) = a \in \{0, 1\}$, and likewise Bob sends back some $X_B(y) = b \in \{0, 1\}$.

(iii) The verifier can then accept the 4-tuple (x, y, a, b) iff we have $a = b \iff x = y$.

Exercise 1. In this scenario, when will the verifier be able to catch the lie, i.e. in what situations will the verifier accept when graph is not actually 2-colorable?

Another possible protocol is as follows: Suppose $X_A(x) = X_B(x) = x \mod 2$. Then, one can check easily that the success probability is $1 - 1/(2n)$.

Exercise 2. Prove that this is the best you can do.

Definition 1. We define a 2-prover game $G = (\mu, V)$ as follows: We have two provers A, B which are given inputs x, y from a verifier, where (x, y) is selected according to the distribution μ supported on $\mathcal{X} \times \mathcal{Y}$ (usually finite sets) as follows. After receiving x, y from the prover, the verifier is sent $a \in \mathcal{A}$ and $b \in \mathcal{B}$ from A, B, respectively. Moreover, $V : \mathcal{X} \times \mathcal{Y} \times \mathcal{A} \times \mathcal{B} \to \{0, 1\}$ is a function that the verifier then uses to either accept or reject, with the former corresponding to 1 and the latter to 0. The goal of the provers is to send back a, b so that $V(x, y, a, b) = 1$.

In our previous case, we had $\mathcal{X} = \mathbb{Z}_n = \mathcal{Y}$. We also had $V : \mathcal{X} \times \mathcal{Y} \times \mathcal{A} \times \mathcal{B} \to \{0, 1\}$. In the previous case we moreover had $\mathcal{A} = \mathcal{B} = \{0, 1\}$.

Definition 2. We define a strategy $f : \mathcal{X} \to A, g : \mathcal{Y} \to B$ as a pair of functions used by the two provers. This gives rise to a notion of a **game value dependent on** f, g, namely

$$\mathrm{val}(G, f, g) = \mathbb{E}_{(x,y)\sim\mu}\left[V(x, y, f(x), g(y)\right]$$

is the value of G given strategy f, g. This moreover gives rise to the **game value**, namely

$$\mathrm{val}(G) = \omega(G) = \max_{f,g} \left\{\mathrm{val}(G, f, g)\right\}.$$

For now, suppose that the strategies of the provers are deterministic.

Given a game, how can we compute its value? Note that this has to be a hard question, because this should be able to answer questions such as whether or not graphs are 3-colorable. In '92, it was known that there was a family of games for which it is hard to approximate values to within $\pm 10^{-10}$. In like manner, are there games whose values are hard to approximate additively to $1 - \varepsilon$. Idea: let's just repeat the same game a bunch of times!

Repetition of Games

We will no longer be in the 2-prover game case. Instead of just one question, we can ask a bunch of questions. It turns out that the analysis is a function of how we ask these questions. In sequential repetition, where we ask questions one after another, we have $\mathrm{val}\left(G^{k\text{-seq-rep}}\right) = \mathrm{val}(G)^k$.

In parallel repetition, we have $\mu^{\otimes k} = k$-fold product of μ and

$$V^{\otimes k}\left((x_1, \ldots, x_k), (y_1, \ldots, y_k), (a_1, \ldots, a_k), (b_1, \ldots, b_k)\right) = \bigwedge_{i=1}^{k} V(x_i, y_i, a_i, b_i).$$

We moreover define $\mathrm{val}\left(V^{\otimes k}, \overline{f}, \overline{g}\right)$, where $\overline{f} : \mathcal{X}^k \to \mathcal{A}^k, \overline{g} : \mathcal{Y}^k \to \mathcal{B}^k$ in the same fashion as earlier. Note that it is not necessarily true that $\overline{f} = f^{\otimes k}$ or even that $\overline{f}$ can be written as a k-fold product.

Definition 3. We define the **game value** as

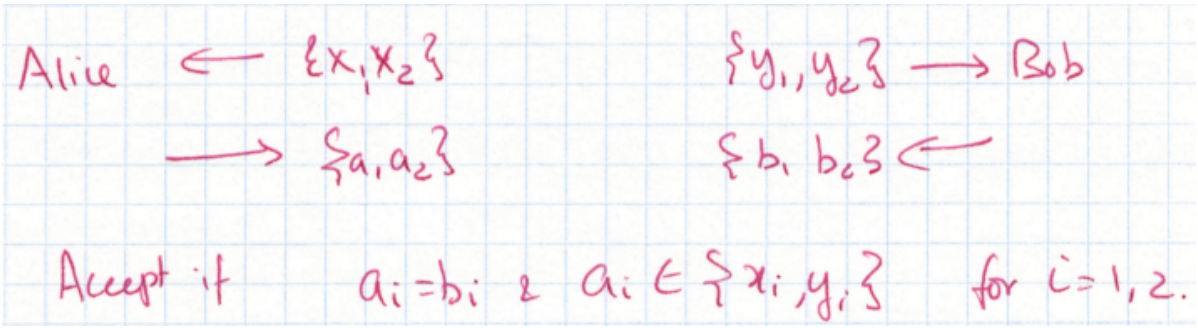
$$\omega(G^{\otimes k}) = \max_{\overline{f},\overline{g}} \left\{\mathrm{val}\left(G^{\otimes k}, \overline{f}, \overline{g}\right)\right\}.$$

Question:

$$\omega\left(G^{\otimes k}\right) = \omega(G)^k?$$

Note that $\geq$ is easy, but $\leq$ is elusive.

Example 2 (Feige's Counterexample)**.** There exists some G with $\omega(G^2) = \omega(G) = 1/2$. In fact, for all s, there are even G_s so that $\omega(G^{\otimes S}) = \omega(G) < 1$. This suggests that perhaps we can repeat games without changing the value. In this example, the verifier tosses two coins.



Clearly, Alice can easily guess x and Bob can easily guess y. Note that probability of success is at most $1/2$; at least one of Alice and Bob will have to guess the other person's coin (since we only accept if both guesses are equal and correct). We can reformulate the game as follows (for instance, guessing 3 means "I guess $y = 3$"). Then, $V(x, y, a, b)$ accepts iff $a = b$ and $b \in \{x, y\}$. It's easy to check that $\omega(G) = 1/2$.

As a general fact, note that $\omega(G^{k-1}) \geq \omega(G^k)$.

Exercise 3. Why is this true?

Remark 1. There is a modification of the game: $\mathcal{A} = \mathcal{B}, \mathcal{X} = \mathcal{Y}, f : \mathcal{X} \to A$ and $V : \mathcal{X} \times \mathcal{X} \times \mathcal{A} \times \mathcal{A} \to \{0, 1\}$ and $\mu \sim \mathcal{X} \times \mathcal{X}$. This is denoted the "PCP" version of the game. We have $\mathrm{val}(G, \mu) = \max_f \mathrm{val}(G, f)$. For this case, apparently, the inequality above may not hold.

Example 3. So clearly we have $\omega(G^{\otimes 2}) \leq 1/2$. In the 2-fold repetition strategy, we hope that $x_1 = y_2 - 2$. If we treat "hope" as an event, the probability of this is $1/2$ (in particular, x_1 is $y_1 - 2$ or its complement with equal probability). Note that $\Pr[\text{win}|\text{hope}] = 1$: the strategy is $f(x_1, x_2) = (x_1, x_1 + 2)$ and $g(y_1, y_2) = (y_2 - 2, y_2)$. In particular, this ensures that Alice and Bob both win together or lose together – we are tying the win/loss together, which should improve the win probability.

Verbitsky answers this question: for all G, with $\omega(G) < 1$, for all $\varepsilon > 0$, there exists $k = k(G, \varepsilon)$ such that $\omega(G^{\otimes k}) < \varepsilon$. Essentially every game eventually shrinks if you repeat it enough.

Then Raz came along and said the following:

Theorem 1. *For all $\mathcal{A}, \mathcal{B}, \varepsilon > 0$, there exists $\delta > 0$ so that for all G and for all k, if $\omega(G) \leq 1 - \varepsilon$, we have $\omega(G^{\otimes k}) = (1 - \delta)^k$.*

Raz's Lemma

Now, suppose $S \subset [k]$ and fix some strategy $\overline{f}, \overline{g}$. Let W_S = event that we win on coordinates $i \in S$. Equivalently, $V(x_i, y_i, \overline{f}(x)_i, \overline{g}(y)_i) = 1$ for every $i \in S$. It would be nice to have $\Pr\left[W_i | W_{\{1,\ldots,i-1\}}\right] \leq \omega(G)$

Exercise 4. Why doesn't this work? Give an example from before.

Precisely, we will use the following lemma in the proof of the main theorem next time:

Lemma 1. *For every $\mathcal{A}, \varepsilon > 0$, there exists some $\gamma > 0$ such that for all G with $\omega(G) \leq 1 - \varepsilon$ for all $\overline{f}, \overline{g}, k$, For every subset $S \subset [k], |S| < \gamma k$, one of the two happens:*

(i) There exists some $i \notin S$ so that $\Pr[W_i | W_S] \leq 1 - \varepsilon/2$.

(ii) $\Pr[W_S] \leq 2^{-\gamma k}$.

Why is this important? Suppose $S_0 = \varnothing$. Note that $W_{S_0} = 1$. There exists some $i_0 \notin S_0$ so that $\Pr[W_{S_0}] \leq 1 - \varepsilon/2$. We can write $S_1 = S_0 \cup \{i_0\}$. We have $\Pr[W_{S_1}] \leq 1 - \varepsilon/2$. Then there is some $i_2 \notin S_1$ so that $\Pr[W_{i_2} | W_{S_1}] \leq 1 - \varepsilon/2$. If $S_2 = S_1 \cup \{i_2\}$, then $\Pr[W_{S_2}] \leq (1 - \varepsilon/2)^2$. The only reason we would stop is if we run into the second condition or run out of elements of $[k]$. In the former, we have $\Pr\left[W_{[k]}\right] \leq \Pr[W_S] \leq 2^{-\gamma k}$, which is exponentially small in k. In most cases, we should expect that $\Pr\left[W_{[k]}\right] \leq (1 - \varepsilon/2)^{\gamma k}$ by repeating the process above.

Note that currently our verifiers have randomness but our provers do not. In the next lecture, we will use correlated sampling and provide randomness to our provers to continue our analysis.