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

Lecture 18

Instructor: Madhu Sudan *Scribe: David Xiang*

1 Outline

1. Introduce 2-Prover games
2. Define and state Raz's Parallel Repetition Theorem
3. Introduce key lemma

The main references for this material will be Raz[1] and Holenstein[2].

2 2-Prover Games

2.1 Definitions

The setting of a 2-Prover game consists of a verifier and two provers A and B, referred to as Alice and Bob respectively. Alice and Bob should be thought of as having some claim based off of insider knowledge that the verifier is attempting to verify, which is modeled as the verifier generating questions (x, y) according to some distribution μ. The verifier will then separately send question x to Alice and question y to Bob, who return answers a and b respectively. The verifier then decides whether to accept or reject the claim.

Definition 1. *A* ***2-prover game*** *G consists of sets $\mathcal{X}, \mathcal{Y}$ (the set of possible questions), $\mathcal{A}, \mathcal{B}$ (the set of possible responses, and a pair (μ, V), where μ is a distribution supported on $\mathcal{X} \times \mathcal{Y}$, and V is a function*

$$\mathcal{X} \times \mathcal{Y} \times \mathcal{A} \times \mathcal{B} \to \{0, 1\}$$

V here represents the decision of the verifier to accept/reject the claim, which depends on the questions asked and the answers received.

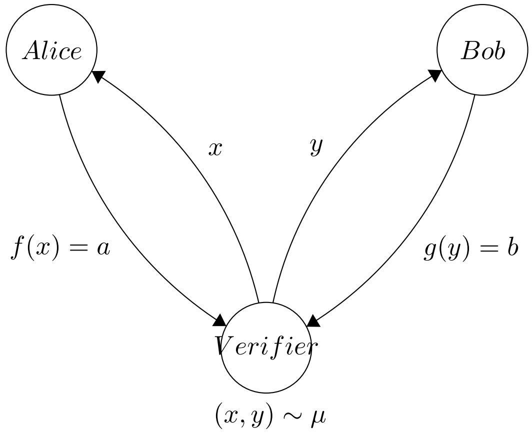
Definition 2. *A* ***strategy*** *is a pair of deterministic functions $f : \mathcal{X} \to \mathcal{A}, g : \mathcal{Y} \to \mathcal{B}$.*

Definition 3. *The value of a game under a strategy $Val(G, f, g)$ is*

$$\mathbb{E}_{(x,y)\sim\mu}[V(x, y, f(x), g(y)]$$

i.e. the probability that the verifier confirms Alice and Bob. The ***value*** *of G, $\omega(G)$, is then*

$$\omega(G) = \max_{f,g}\{\, Val(G, f, g)\}$$



2.2 Odd Cycle

Our first example of a game will consist of graph coloring. The general setting consists of a large public graph G, which Alice and Bob claim is c-colorable. The verifier would like to verify this fact, but is unwilling to actually attempt to find a c-coloring for G (for example, G could be very large) and so will instead ask Alice and Bob a question to convince itself that the graph is c-colorable.

For simplicity we'll consider the case where G is an odd cycle and $c = 2$. To be exact, we'll take G to be the graph with vertex set indexed by elements of $\mathbb{Z}_n$, n odd, and with edges between vertices i to $i + 1$ for all $i \in \mathbb{Z}_n$. Such a graph is clearly not 2-colorable, and so ideally the verifier will be able to identify whether Alice and Bob are lying.

A first try at a verification setup may be to randomly choose i, and then ask Alice and Bob for the colors of the vertices i and $i + 1$ in a 2-coloring of the graph G. The problem here is that Alice can return 0 and Bob can return 1 (irrespective of the query), and the verifier will not be able to detect any lie from this. A slightly better strategy will be as follows:

1. Uniformly pick i in $\mathbb{Z}_n$, and then generate the pair $(x, y) = (i, i)$ with probability $\frac{1}{2}$ and $(x, y) = (i, i+1)$ otherwise.

2. Ask Alice for the color of vertex x and Bob for the color y

3. Alice and Bob return colors χ_A and χ_B respectively. We accept $(V(x, y, \chi_A, \chi_B) = 1)$ if their answer is consistent with a 2-coloring of the graph, and reject $(V(x, y, \chi_A, \chi_B) = 0)$ otherwise.

Suppose Alice and Bob agree to return $x \pmod 2$ and $y \pmod 2$ respectively. We will only ever detect a problem if $x = n - 1, y = n$, and so under this strategy Alice and Bob succeed with probability $1 - \frac{1}{2n}$.

Exercise 4. *Show this strategy is optimal, in the sense that the value of this game is* $1 - \frac{1}{2n}$.

Proof. The only way for Alice and Bob to do better than a value of $1 - \frac{1}{2n}$ is to win with probability 1. Assume Alice and Bob have strategies which do so- in this case, since we generate pairs $(x, y) = (i, i)$ with nonzero probability for all i, it follows that $f(i) = g(i)$ for all i. But now if Alice and Bob are to win with probability 1, they also need $f(i)! = f(i + 1)$ for all i (with indices taken $mod n$), which implies that $f(i) = f(i + 2k)$ for all integer k, but taking $k = \frac{n+1}{2}$ shows that $f(i) = f(i + 1)$, a contradiction. □

3 Repetition of games

3.1 Parallel and Sequential Repetition

It's a pretty natural question to ask what the value of a game is. However, as the previous example alludes to, difficult problems such as 3-coloring can be encoded into a game, and so asking for the value of a game is as least as hard as asking questions such as graph colorability, as we can construct games which have value 1 iff a graph is c-colorable. In fact, we have the even stronger result

Theorem 5. *There exists an infinite family of games whose value is hard (NP-complete) to approximate to within* $\pm 10^{-10}$.

Does this imply that games whose values are hard to approximate additively to within $1-\varepsilon$, for arbitrary ε? There's in general a canonical procedure to go from these sorts of weak approximation results to strong results through amplification- run the game many times in parallel, to get a game with much lesser value. This sort of thinking motivates the parallel repetition theorem. It's important to first define what we mean by repetition, however.

1. We can repeat the game multiple times, so that each game is independent of each other. This is **sequential repetition**, and a game repeated k times in this fashion will be denoted G^k.

2. We can generate k questions and give them to Alice and Bob at once (think of this as a long multi-part question). This is called **parallel repetition**, and is the more mathematically interesting of the two, since Alice's (resp. Bob's) answer to a particular question can now depend on their answers to previous questions. Under our definitions this is still a 2-prover game, unlike in the case of sequential repetition. A game repeated k times in this fashion will be denoted $G^{\otimes k}$.

Exercise 6. *Show that* $\omega(G^k) = \omega(G)^k$.

Proof. Since each query is independently satisfied with probability at most $\omega(G)$, the probability all queries are satisfied is $\omega(G)^k$. □

Exercise 6 is confirmation of the claim that parallel repetition is the more interesting of the two types of repetition. Let's be formal about the definitions now:

Definition 7. *Given a game* $G = (\mu, V)$, *we define the k-fold parallel repetition of G to be* $G^{\otimes k} = (\mu^{\otimes k}, V^{\otimes k})$, *where* $\mu^{\otimes k}$ *is the k-fold product distribution and* $V^{\otimes k}$ *is the map*

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

The important aspects here are that the strategies $\overline{f}, \overline{g}$ are now functions from $\mathcal{X}^k \to \mathcal{A}^k, \mathcal{Y}^k \to \mathcal{B}^k$ respectively. There is no such assumption that $\overline{f}(x_1, \dots x_k) = (f(x_1), f(x_2), \dots f(x_k))$, which is essentially what is going on in the case of sequential repetition.

We can ask again whether

$$\omega(G^{\otimes k}) = \omega(G)^k,$$

the bound $\omega(G^{\otimes k}) \geq \omega(G)^k$ follows from just treating parallel repetition as sequential repetition.

3.2 Feige's Counterexample

In fact, for all k there exists G with $\omega(G^{\otimes k}) = \omega(G) < 1$, and so the previous claim cannot hold. We'll show this in the case $k = 2$, this counterexample is due to Feige. Define a game G as follows:

1. The verifier tosses two fair coins $x \in \{1, 2\}, y \in \{3, 4\}$. (i.e. x takes the values in $\{1, 2\}$ with equal probability, likewise for y) The verifier sends x to Alice and y to Bob

2. Alice and Bob try to guess the value of one of the coins they receieved; that is, Alice and Bob both try to guess the value of either x or y.

3. The verifier accepts if both guesses are equal and correct. In our notation, we accept if $a = b$ and $b \in \{x, y\}$

Alice can easily guess the value of x, and Bob can easily guess the value of y. But they need to guess the value of the same coin- and it's clear from this that intuitively the value of the game is $\frac{1}{2}$, since WLOG Alice has to guess Bob's coin (it's not too hard to enumerate over all possible functions to see that the value of the game is $\frac{1}{2}$).
As an aside we should note that even the fact that $\omega(G^{\otimes 2}) \geq \omega(G)$ is very subtle - for example, if we change the setting to $\mathcal{X} = \mathcal{Y}$, so that $V : \mathcal{X} \times \mathcal{X} \times \mathcal{A} \times \mathcal{A} \to \{0, 1\}$, with $Val(G) = \max_f \text{val}(G, f)$, it's possible to have $\omega(G^{\otimes k}) > \omega(G)$. (This setting is known as the "PCP version" of the game. See page 18 of [3] for a more detailed example)

In the parallel version of the game, Alice and Bob can now leverage the fact that Alice and Bob know the values of x and y. Let's denote the values of the coin flips by x_1, x_2, y_1, y_2. The idea is to hope that the event $E : x_1 = y_2 - 2$ happens. This happens with probability $\frac{1}{2}$. In this case, Alice and Bob can try to guess the values of x_1 and y_2: Alice will guess $(x_1, x_1 + 2)$, and Bob will guess $(y_2 - 2, y_2)$. By construction Alice and Bob win iff E occurs, and so the value of this game is $\frac{1}{2}$, demonstrating that

$$\omega(G^{\otimes 2}) = \omega(G) = \frac{1}{2}$$

in this case.

4 Parallel Repetition Theorem

4.1 Main Theorem

We just saw that it can be the case that $\omega(G^{\otimes k}) = \omega(G)$- which is not very useful for our original motivation of being able to apply amplification to these types of problems. Thankfully, this is not the case in general.

Theorem 8 (Verbitsky)**.** *For every game G with $\omega(G) < 1$ and for all $\varepsilon > 0$ there exists k such that $\omega(G^{\otimes k}) < \varepsilon$.*

This shows the intuitive claim that parallel repetition should eventually make games harder. The problem here is that k is a function of both ε and G, and so isn't very useful for our purposes. Raz's parallel repetition theorem is able to remove the dependence on G.

Theorem 9 (Raz[1])**.** *For all answer sets $\mathcal{A}, \mathcal{B}$ and $\varepsilon > 0$, there exists $\delta > 0$ such that for all G, k then*

$$\omega(G) < 1 - \varepsilon \Rightarrow \omega(G^{\otimes k}) = (1 - \delta)^k$$

The important part here is that δ is a function only of the size of $\mathcal{A}, \mathcal{B}$ (which are usually constant) and ε. We will work towards the proof of this result in the next few lectures.

4.2 A useful lemma

For fixed $\overline{f}, \overline{g}$ and $S \subset [k]$, let w_S be the event that the provers win on all coordinates $i \in S$. If $\mathbb{P}[w_i | w_{1,\ldots i-1}] \leq w(G)$ we'd be able to conclude that $\omega(G^{\otimes k}) \leq \omega(G)^k$, except we know from Feige's example that this does not hold: in that case $P(w_2 | w_1) = 1$. What we can say is

Lemma 10. *For all $\mathcal{A}, \mathcal{B}, \varepsilon > 0$, there exists $\gamma > 0$ such that for all $G, \overline{f}, \overline{g}, k$ with $\omega(G) \leq 1 - \varepsilon$, the following holds: For any subset S of $[k]$ with size $|S| < \gamma k$, there exists $i \notin S$ so that either*

$$\mathbb{P}[w_i | w_S] \leq 1 - \frac{\varepsilon}{2}$$

or

$$\mathbb{P}[w_S] \leq 2^{-\gamma k}.$$

Why is this useful? We can start with

$$S_0 = \emptyset \Rightarrow w_{S_0} = 1$$

Then, we can find $i_1 \notin S_0$ with $\mathbb{P}(w_{i_1}) \leq 1 - \frac{\varepsilon}{2}$, and we can continue inductively to build our sets S_i. At any step $P(w_{S_j}) \leq (1 - \varepsilon/2)^j$.
We only stop when $w_S \leq 2^{-\gamma k}$, but this bounds $P[w_{[k]}] \leq P[w_S] \leq 2^{-\gamma k}$, which is exponentially small in k. Or, we never stopped, and so instead we have the bound $P(w_{[k]}) \leq (1 - \varepsilon/2)^{\gamma k}$, the important conclusion is that both bounds are exponentially small in k.
How would we go to prove a lemma like this? The strategy will be to use a simulation argument- sneak in the ith coordinate into a k-fold parallel repetition game, and calculate the probability of winning from there.

References

[1] Raz, R. (1998). A Parallel Repetition Theorem. SIAM Journal on Computing, 27(3), 763-803.

[2] Holenstein, T. (2007). Parallel repetition: Simplifications and the no-signaling case. Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing, 411-419.

[3] Radhakrishnan, J. and Sudan, M. (2006). On Dinur's proof of the PCP Theorem. Bulletin (New Series) of the American Mathematical Society, Volume 44, Number 1, January 2007, Pages 19–61 S 0273-0979(06)01143-8