COS597D: Information Theory in Computer Science November 9, 2011

Lecture 16

Lecturer: Mark Braverman *Scribe: Gabriel Cadamuro**

1 Comparison between private and public randomness

In the previous lecture we defined three ways to evaluate the cost of a protocol: deterministically, using public randomness and using private randomness. Clearly $D(f) \geq R_\epsilon^{pri}(f) \geq R_\epsilon^{pub}(f)$ for any ϵ. However it was also hinted at that private and public randomness are very close to one another. This would be a good situation as public randomness is easier to analyze and so we could restrict our attention to that model.

We introduce first an inequality that we do not prove but has several applications in probability and will be used later.

Theorem 1. *Hoeffding's inequality. Let $x_1...x_t$ be i.i.d on [0,1]. Denote the empirical mean $\frac{\sum_{i=1}^{t} x_i}{n}$ by $\bar{x}$. Then $Pr[\bar{x} - E(\bar{x}) \geq \delta] \leq \exp(-2\delta^2 t)$.*

Theorem 2 (Newman'91)**.** *Let $f: X \times Y \mapsto Z$ as in lecture 13. Then for each $\delta, \epsilon > 0$*

$$R_{\epsilon+\delta}^{pri}(f) \leq R_\epsilon^{pub}(f) + O(log(n) + log(\frac{1}{\delta}))$$

Proof Let us start with the public protocol. Given some random coin flip result r let $P(x, y, r)$ be the answer returned by the protocol with inputs x and y. Let $Z(x, y, r)$ then be 1 if $P(x, y, r) \neq f(x, y)$ and 0 otherwise. Clearly, $\forall x, y \in X, Y : Pr[Z(x, y, z) = 1] \leq \epsilon$, where the probability is taken over r.
Let us now define a private protocol. Before the protocol starts, a (fixed) random set of strings $\{r_i\}_{i=1}^t$ is selected. During the execution of the protocol, pick random strings $\{r_i\}$ using private randomness. We will now bound $Pr[\sum_{i=1}^{t} Z(x, y, r_i) \geq (\epsilon + \delta)t]$ using the Hoeffding inequality mentioned beforehand. Picking $t = \lceil \frac{n}{\delta^2} \rceil + 1$ gives us $Pr[\bar{Z} - E[\bar{Z}] \geq \delta] \leq \exp(-2\delta^2 \lceil \frac{n}{\delta^2} \rceil + 1)$ which reduces to $Pr[\bar{Z}t \geq (\delta + \epsilon)t] \leq \exp(-2n)$. So what we have shown is that t randomly selected values of r will lead to a protocol with error $> \epsilon + \delta$ on (x, y) less than $\exp(-2n)$ percent of the time. This then allows us to use the union bound to show there exists a set of t r_i's that has $Pr[\bar{Z}t \geq (\delta + \epsilon)t]$ for every single combination of (x, y). Hence running the public protocol with the random bits picked from one of $\{r_i\}$ is a valid private-coin protocol with error $\epsilon + \delta$. Finally we note that this requires us to transmit which one of the t values Alice will use to Bob which takes $log(t) = log(n) + 2log(\frac{1}{\delta})$ bits in addition to the same number of bits as the public protocol would, hence obtaining the bound of the theorem. ■

2 Distributional complexity

Up to now we have only considered randomness that is not dependent on the input. Namely, given any input we are guaranteed to return the correct results with probability $1 - \epsilon$. However, it may be the case that we

*Some use was made of "Communication Complexity" by Kushilevitz and Nisan. All logarithms are taken with base 2.

know something about the type of inputs that will be sent and are certain that inputs are drawn from some distribution μ. The question then would be whether we can get the same error bound for smaller cost. This leads to distributional complexity.

Definition 3. *Let μ be a distribution on $X \times Y$ and ϵ be an error margin. Then we define $D^{\mu}_{\epsilon}(f)$ to be the cost of the smallest deterministic protocol that computes f correctly $1-\epsilon$ fraction of the time given the distribution μ on the inputs.*

The replacement of the possibility of an adversarial selection of inputs with a totally known distribution can yield counter-intuitive results.

Example 4. *If f is the equality function EQ then $D^{uniform}_{0.25}(EQ) = 0$. This is the case as we can simply return 0 each time, confident in the fact that we will return an error only 2^{-n} fraction of the time.*

We now begin a series of examples that keeps the same function but varies the input distribution μ. We will get very different costs for each distribution.

Example 5. *Now let us try to solve the disjoint function DISJ where the distribution $\mu = A, A^c$ with probability $\frac{1}{2}$ and $\mu = A, A^c \cup \{a\}$ with probability $\frac{1}{2}$ $(a \in A)$. Now we claim $D^{\mu}_{0.25}(DISJ) \leq 2$. This is Alice can simply send the parity of the number of 1's in her input over to Bob. Bob can then see whether his input has parity consistent with the complement of something with that parity and return 1 or 0 appropriately.*

Example 6. *We take the same function but now μ is such that any index in X or Y has value 1 with probability $\frac{1}{\sqrt{n}}$, and the indices of X and Y are drawn independently. We claim $D^{\mu}_{0.1}(DISJ) \leq O(\sqrt{n}log(n))$.If the protocol consists of Alice sending over the first $O(\sqrt{n})$ indices where her input equals 1 then we note this costs $log(n)$ bits per index and there are $O(\sqrt{n})$ such indices to transmit, proving the size assertion. This bound turns out to be essentially tight.*

Example 7. *Finally, we let μ be such that we draw each value pair for an index of x, y over the following cases $x_i = 0, y_i = 0$; $x_i = 0, y_i = 1$ and $x_i = 1, y_i = 0$ with probability $\frac{1}{3}$ each and $x_i = 1, y_i = 1$ for a randomly selected i with probability $\frac{1}{2}$. For sufficiently small ϵ the cost becomes $\Omega(n)$.*

Now note that $D^{\mu}_{\epsilon} \leq R^{pub}_{\epsilon}(f)$, as any protocol which gets $1-\epsilon$ confidence when inputs can be chosen by an adversary will certainly get the same confidence on a fixed probabilistic distribution on inputs. As with examples 4 and 5, we can also pick simple μ for which $D^{\mu}_{\epsilon}(f) < R^{pub}_{\epsilon}(f)$. However in example 7 we obtained a μ such that $D^{\mu}_{\epsilon} = \Theta(R^{pub}_{\epsilon}(f))$ which leads us to wonder how D^{μ}_{ϵ} and $R^{pub}_{\epsilon}(f)$ are related. Surprisingly, there is a very neat answer.

Theorem 8. $R^{pub}_{\epsilon}(f) = \max_{\mu}(D^{\mu}_{\epsilon}(f))$

Essentially this states that there are pathological distributions which make it as costly to solve the function as not knowing about the distribution in advance at all!
In order to better analyze this problem, we will introduce some notions from game theory that apply to this case.

Definition 9. *A zero sum game for two players is a game where the rewards for players 1 and 2 sum to 0 regardless of the end state. The payoff of the game $p_1(s_1, s_2)$ for player 1 is a function of the strategies s_1 and s_2 chosen by the players. By definition the payoff of the game for player 2 is given by $p_2(s_1, s_2) = -p_1(s_1, s_2)$. A mixed strategy is a probability distribution over strategies. The payoff of a mixed strategy (or a pair of mixed strategies for the two players) is the expected payoff of the game under the mixed strategy.*

Theorem 10. *The min-max theorem. Let players 1 and 2 be playing a zero sum game. $\exists$ a value v and mixed strategies for players 1 and 2 S_1, S_2 such that*
(1) If player 2 plays with S_2 then player 1's best payoff is v
(2) If player 1 plays with S_1 then player 2's best payoff is -v.

In other words, the first condition guarantees that player 2 can secure a payoff of $-v$, while the second condition guarantees that player 1 can secure a payoff of v – thus making v the "correct" payoff for player 1. With these tools we can now prove theorem 8.

Proof (Theorem 8) We want to show $R^{pub}_{\epsilon}(f) \leq \max_{\mu}(D^{\mu}_{\epsilon}(f))$, as the other direction is simple and has already been discussed.
Let us start by fixing f and ϵ. Let $c = \max(D^{\mu}_{\epsilon}(f))$. Now let player 1 be the player who can pick a deterministic protocol π of size $\leq c$ and player 2 the player who can pick any input $(x, y) \in X \times Y$. Let the payoff for player 1 be
$Payoff_1 = 1$ if $\pi(x, y) = f(x, y)$; 0 otherwise.
$Payoff_2 = -Payoff_1$, so we have a zero sum game.
Noting that $c \geq D^{\mu}_{\epsilon}(f)$ for any (x,y), we see that for any given mixed strategy made by player 2 (i.e an input distribution μ) player 1 has a strategy which gives him payoff $\geq 1 - \epsilon$ (simply using the protocol designed for the distribution). However, the min-max theorem notes that this implies player 1 has a mixed strategy S' which gives the same payoff ($\geq 1 - \epsilon$) for any mixed strategy (i.e any input distribution μ) that player 2 gives.
Now let us consider this mixed strategy S'. This is a distribution on many deterministic c-bit protocols $\{\pi_i\}$ which has error $< \epsilon$ for any input distribution. However, this would be equivalent to looking at some random coin flips, picking a c-bit π_i based on that, and then running π_i on the input. However, this means S' is precisely a publicly random protocol that has error $< \epsilon$ and size $\leq c$. This proves $R^{pub}_{\epsilon}(f) \leq \max_{\mu}(D^{\mu}_{\epsilon}(f))$.
■

3 Average Complexity

We now sketch another way to consider the communication complexity of a function that moves away from asking about worst case cost and more towards the average.
Let $c(\pi, x, y)$ be the number of bits that protocol π exchanges when given inputs x and y. For a random protocol define $c(\pi, x, y, r)$ similarly except that it also depends on the random string result r. Recalling that $D^{\mu}_{\epsilon}(f) = \min_{\pi}(\max_{x,y}(c(\pi, x, y)))$ over all protocols π that solve f on μ with error at most ϵ we get the following:

Definition 11. $\bar{D}^{\mu}_{\epsilon}(f) = \min_{\pi}(E[c(\pi, x, y)])$ *where the expectation is taken over the inputs x,y being drawn from* μ.

Definition 12. $\bar{R}^{pub}_{\epsilon}(f) = \min_{\pi}(\max_{x,y}(E[c(\pi, x, y, r)]))$ *where the expectation is over the possible different random coin flip results.*

We now given an example and state, but do not prove, a lemma which links distributional complexity and public randomness complexity in this model.

Example 13. $\bar{D}^{uniform}_{0}(EQ) = O(1)$. *Note that the chance of two independently and uniformly sampled inputs being equal at any index is* $\frac{1}{2}$, *therefore the expected number of bits that will be transmitted before a discrepancy is* $\sum_{i=1}^{n} 2^{-i} i$ *which is indeed* $O(1)$.

Lemma 14. $\frac{1}{2} \max_{\mu} \bar{D}^{\mu}_{\frac{\epsilon}{2}}(f) \leq \bar{R}^{pub}_{\epsilon}(f) \leq 2 \max_{\mu} \bar{D}^{\mu}_{2\epsilon}(f)$

The proof for this lemma proceeds in a similar way to that for theorem 8 by using the minmax theorem for zero sum games.