COMPUTATIONAL ENTROPY

HANI T. DAWOUD

Pseudorandom Generators. For a class $\mathcal{C}$ of functions, find a distribution D such that

(1) D fools $\mathcal{C}$: $\forall f \in \mathcal{C}, f(D) \approx f(U)$.
(2) D is efficiently samplable.
(3) D is sampled using a few random bits.

For (1) and (2): D can be the uniform distribution. For (1) and (3): $\forall \mathcal{C}$ $\exists$ inefficiently samplable D using $O(\log\log|\mathcal{C}|)$ random bits. How? Use the *probabilistic method.* A random function is PRG with high probability.

Notation. $X \equiv^c U_n$ denotes that X and Y are computationally indistinguishable. U_k is a uniform random variable over $\{0,1\}^k$.

A random variable X over $\{0,1\}^n$ is *pseudorandom* if $X \equiv^c U_n$. As a sequence of bits, $X = (X_1, X_2, \ldots, X_n)$ is *unpredictable* if $\forall i \in [n]$ and $\forall$ PPT P, $\mathbb{P}\left[P(X_1, X_2, \ldots, X_{i-1}) = X_i\right] \leq \frac{1}{2} + \mathsf{negl}$.

Theorem 0.1 (Pseudorandomness vs. Unpredictability)**.** *X is pseudorandom iff it is unpredictable. Or, $X \equiv^c U_n$ iff $(X_1, X_2, \ldots, X_{i-1}, X_i) \equiv^c (X_1, X_2, \ldots, X_{i-1}, U_1) \forall i \in [n]$.*

The notion of pseudorandomness can be generalized. X has *pseudoentropy* at least k if $\exists Y$ with $H(Y) \geq k$ such that $X \equiv^n Y$. X has *pseudo-min-entropy* at least k if $\exists Y$ with $H_\infty(Y) \geq k$ such that $X \equiv^n Y$. (In the special case when $k = n$, X is pseudorandom.) Pseudoentropy and pseudo-min-entropy can be generalized even further. Let (X, B) be jointly distributed. B has conditional pseudoentropy at least k given X if $\exists C$ jointly distributed with X with $H(C|X) \geq k$ such that $(X, B) \equiv^c (X, C)$. B has conditional pseudo-min-entropy at least k given X if $\exists C$ jointly distributed with X with $H_\infty(C|X) \geq k$ such that $(X, B) \equiv^c (X, C)$.

The remark below asserts the case when the pseudoentropy notion is interesting.

Remark 0.2. *By definition, any random variable X has pseudoentropy at least $H(X)$: Take Y to be an independent random variable distributed identically to X. The interesting case is when the pseudoentropy of a random variable is strictly greater than its real entropy. For example, the pseudoentropy of $X \sim U_n$ is n and $H(X) = n$. But, $G(U_n)$ for a PRG $G : \{0,1\}^n \to \{0,1\}^m$ has pseudoentropy $m > n$ by definition, while $H(G(U_n)) \leq n$. The difference between the pseudoentropy of a random variable X and its real entropy is the "entropy gap", defined as $\Delta = \text{pseudoentropy of } X - H(X)$. Thus the notion of pseudoentropy is only interesting when $\Delta > 0$.*

Next is a remark on how to interpret the notion of conditional pseudoentropy.

Remark 0.3. *If B has pseudoentropy at least k given X, then (X, B) has pseudoentropy at least $H(X) + k$. (The pseudoentropy of X, which is at least $H(X)$, plus the pseudoentropy of B given X, which is at least k.) The converse is not necessarily true. Consider X having pseudoentropy at least $H(X) + k$ on its own, and B being completely determined by X.*

The notion of unpredictability can also be generalized.

1. Entropy

For a random variable X and $x \in \mathrm{Supp}(X)$, the *sample-entropy* of x with respect to X is

$$\mathrm{H}_X(x) \triangleq \log\left(\frac{1}{\mathbb{P}\left[X = x\right]}\right).$$

The *Shannon entropy* is

$$\mathrm{H}(X) \triangleq \mathop{\mathbb{E}}_{x \sim X}\left[\mathrm{H}_X(x)\right].$$

Date: July 18, 2016.

The *min-entropy* is

$$\mathrm{H}_\infty(X) \triangleq \min_{x \in \mathrm{Supp}(X)} \mathrm{H}_X(x).$$