Chapter 27

Even more on Entropy, Randomness, and Information

By Sariel Har-Peled, December 7, 2009[①]
Version: 0.1

> "It had been that way even before, when for years at a time he had not seen blue sky, and each second of those years could have been his last. But it did not benefit an Assualtman to think about death. Though on the other hand you had to think a lot about possible defeats. Gorbovsky had once said that death is worse than any defeat, even the most shattering. Defeat was always really only an accident, a setback which you could surmount. You had to surmount it. Only the dead couldn't fight on."
>
> – – Defeat, Arkady and Boris Strugatsky

27.1 Extracting randomness

27.1.1 Enumerating binary strings with j ones

Consider a binary string of length n with j ones. $S(n, j)$ denote the set of all such binary strings. There are $\binom{n}{j}$ such strings. For the following, we need an algorithm that given a string U of n bits with j ones, maps it into a number in the range $0, \ldots, \binom{n}{j} - 1$.

To this end, consider the full binary tree $\mathcal{T}$ of height n. Each leaf, encodes a string of length n, and mark each leaf that encodes a string of $S(n, j)$. Consider a node v in the tree, that is of height k; namely, the path π_v from the root of $\mathcal{T}$ to v is of length k. Furthermore, assume there are m ones written on the path π_v. Clearly, any leaf in the subtree of v that is in $S(n, j)$ is created by selecting $j - m$ ones in the remaining $n - k$ positions. The number of possibilities to do so is $\binom{n-k}{j-m}$. Namely, given a node v in this tree $\mathcal{T}$, we can quickly compute the number of elements of $S(n, j)$ stored in this subtree.

As such, let traverse $\mathcal{T}$ using a standard **DFS** algorithm, which would always first visit the '0' child before the '1' child, and use it to enumerate the marked leaves. Now, given a string x of S_j, we would like to compute what number would be assigned to by the above **DFS** procedure. The

key observation is that calls made by the **DFS** on nodes that are not on the path, can be skipped by just computing directly how many marked leaves are there in the subtrees on this nodes (and this we can do using the above formula). As such, we can compute the number assigned to x in linear time.

The cool thing about this procedure, is that we do not need $\mathcal{T}$ to carry it out. We can think about $\mathcal{T}$ as being a virtual tree.

Formally, given a string x made out of n bits, with j ones, we can in $O(n)$ time map it to an integer in the range $0, \ldots, \binom{n}{j}-1$, and this mapping is one-to-one. Let **EnumBinomCoeffAlg**denote this procedure.

27.1.2 Extracting randomness

Theorem 27.1.1 *Consider a coin that comes up heads with probability $p > 1/2$. For any constant $\delta > 0$ and for n sufficiently large:*

1. *One can extract, from an input of a sequence of n flips, an output sequence of $(1-\delta)n\mathbb{H}(p)$ (unbiased) independent random bits.*
2. *One can not extract more than $n\mathbb{H}(p)$ bits from such a sequence.*

Proof: There are $\binom{n}{j}$ input sequences with exactly j heads, and each has probability $p^j(1-p)^{n-j}$. We map this sequence to the corresponding number in the set $S_j = \left\{0, \ldots, \binom{n}{j} - 1\right\}$. Note, that this, conditional distribution on j, is uniform on this set, and we can apply the extraction algorithm of Theorem 27.2.3 to S_j. Let Z be the random variables which is the number of heads in the input, and let B be the number of random bits extracted. We have

$$\mathbf{E}[B] = \sum_{k=0}^{n} \mathbf{Pr}[Z = k]\, \mathbf{E}\Big[B \;\Big|\; Z = k\Big],$$

and by Theorem 27.2.3, we have $\mathbf{E}\Big[B \;\Big|\; Z = k\Big] \geq \left\lfloor \lg \binom{n}{k} \right\rfloor - 1$. Let $\varepsilon < p - 1/2$ be a constant to be determined shortly. For $n(p-\varepsilon) \leq k \leq n(p+\varepsilon)$, we have

$$\binom{n}{k} \geq \binom{n}{\lfloor n(p+\varepsilon) \rfloor} \geq \frac{2^{n\mathbb{H}(p+\varepsilon)}}{n+1},$$

by Corollary 27.2.2 (iii). We have

$$\begin{aligned}
\mathbf{E}[B] &\geq \sum_{k=\lfloor n(p-\varepsilon)\rfloor}^{\lceil n(p+\varepsilon)\rceil} \mathbf{Pr}[Z = k]\, \mathbf{E}\Big[B \;\Big|\; Z = k\Big] \geq \sum_{k=\lfloor n(p-\varepsilon)\rfloor}^{\lceil n(p+\varepsilon)\rceil} \mathbf{Pr}\Big[Z = k\Big]\left(\left\lfloor \lg \binom{n}{k} \right\rfloor - 1\right) \\
&\geq \sum_{k=\lfloor n(p-\varepsilon)\rfloor}^{\lceil n(p+\varepsilon)\rceil} \mathbf{Pr}[Z = k]\left(\lg \frac{2^{n\mathbb{H}(p+\varepsilon)}}{n+1} - 2\right) \\
&= (n\mathbb{H}(p+\varepsilon) - \lg(n+1) - 2)\, \mathbf{Pr}[|Z - np| \leq \varepsilon n] \\
&\geq (n\mathbb{H}(p+\varepsilon) - \lg(n+1) - 2)\left(1 - 2\exp\left(-\frac{n\varepsilon^2}{4p}\right)\right),
\end{aligned}$$

since $\mu = \mathbf{E}[Z] = np$ and $\mathbf{Pr}\left[|Z - np| \geq \frac{\varepsilon}{p} pn\right] \leq 2\exp\left(-\frac{np}{4}\left(\frac{\varepsilon}{p}\right)^2\right) = 2\exp\left(-\frac{n\varepsilon^2}{4p}\right)$, by the Chernoff inequality. In particular, fix $\varepsilon > 0$, such that $\mathbb{H}(p+\varepsilon) > (1-\delta/4)\mathbb{H}(p)$, and since p is fixed $n\mathbb{H}(p) = \Omega(n)$, in particular, for n sufficiently large, we have $-\lg(n+1) \geq -\frac{\delta}{10} n\mathbb{H}(p)$. Also, for n sufficiently large, we have $2\exp\left(-\frac{n\varepsilon^2}{4p}\right) \leq \frac{\delta}{10}$. Putting it together, we have that for n large enough, we have

$$\mathbf{E}[B] \geq \left(1 - \frac{\delta}{4} - \frac{\delta}{10}\right) n\mathbb{H}(p)\left(1 - \frac{\delta}{10}\right) \geq (1-\delta)\, n\mathbb{H}(p),$$

as claimed.

As for the upper bound, observe that if an input sequence x has probability $\mathbf{Pr}[X = x]$, then the output sequence $y = \mathbf{Ext}(x)$ has probability to be generated which is at least $\mathbf{Pr}[X = x]$. Now, all sequences of length $|y|$ have equal probability to be generated. Thus, we have the following (trivial) inequality

$$2^{|\mathbf{Ext}(x)|}\, \mathbf{Pr}[X = x] \leq 2^{|\mathbf{Ext}(x)|}\, \mathbf{Pr}[y = \mathbf{Ext}(x)] \leq 1,$$

implying that $|\mathbf{Ext}(x)| \leq \lg(1/\mathbf{Pr}[X = x])$. Thus,

$$\mathbf{E}[B] = \sum_x \mathbf{Pr}[X = x]\, |\mathbf{Ext}(x)| \leq \sum_x \mathbf{Pr}[X = x] \lg \frac{1}{\mathbf{Pr}[X = x]} = \mathbb{H}(X).$$

■

27.2 From previous lectures

Lemma 27.2.1 *Suppose that nq is integer in the range $[0, n]$. Then* $\dfrac{2^{n\mathbb{H}(q)}}{n+1} \leq \binom{n}{nq} \leq 2^{n\mathbb{H}(q)}$.

Lemma 27.2.1 can be extended to handle non-integer values of q. This is straightforward, and we omit the easy details.

Corollary 27.2.2 *We have:*
(i) $q \in [0, 1/2] \Rightarrow \binom{n}{\lfloor nq \rfloor} \leq 2^{n\mathbb{H}(q)}$. *(ii)* $q \in [1/2, 1]$ $\binom{n}{\lceil nq \rceil} \leq 2^{n\mathbb{H}(q)}$.
(iii) $q \in [1/2, 1] \Rightarrow \frac{2^{n\mathbb{H}(q)}}{n+1} \leq \binom{n}{\lfloor nq \rfloor}$. *(iv)* $q \in [0, 1/2] \Rightarrow \frac{2^{n\mathbb{H}(q)}}{n+1} \leq \binom{n}{\lceil nq \rceil}$.

Theorem 27.2.3 *Suppose that the value of a random variable X is chosen uniformly at random from the integers $\{0, \ldots, m-1\}$. Then there is an extraction function for X that outputs on average at least $\lfloor \lg m \rfloor - 1 = \lfloor \mathbb{H}(X) \rfloor - 1$ independent and unbiased bits.*

27.3 Bibliographical Notes

The presentation here follows [MU05, Sec. 9.1-Sec 9.3].

Bibliography

[MU05] M. Mitzenmacher and U. Upfal. *Probability and Computing – randomized algorithms and probabilistic analysis*. Cambridge, 2005.