

An Ensemble Indistinguishable from Uniform

Proposition 3.2.3 of the textbook claims the existence of an ensemble $X = \{X_n\}_{n \in \mathbb{N}}$ that is polynomial-time-indistinguishable from the uniform ensemble $U = \{U_n\}_{n \in \mathbb{N}}$ but not statistically close to it. The distribution X_n constructed to prove this theorem is uniform on a subset $S_n \subseteq \{0, 1\}^n$ of size $2^{n/2}$.

The proof in the book supplies the low-level details needed to establish this theorem, but it is a little unclear about the construction itself, particularly about how the set S_n is chosen. For the sequel, let $N = 2^{n/2}$ and let $\mathcal{S}_n = \{S \subseteq \{0, 1\}^n \mid \#(S) = 2^N\}$.

The general outline of the proof is as follows. First, we consider an n -input circuit C with at most $2^{n/8}$ gates and we ask about its behavior on uniformly chosen inputs. Let p_C be its expected output. This means that for $p_C \cdot 2^n$ strings x of length n , $C(x) = 1$ and for the other $(1 - p_C) \cdot 2^n$ strings x of length n , $C(x) = 0$.

Next, we define a function $f_C : \mathcal{S}_n \rightarrow \{0, 1\}$, where

$$f_C(S) = \left| \frac{\sum_{s \in S} C(s)}{N} - p_C \right|.$$

Thus, $f_C(S)$ is the deviation of the average value of $C(s)$ over the set S from the average value of $C(u)$ taken over all length- n strings u . Call a set S *bad for C* if $f_C(S) \geq 2^{-n/8}$. Using the Chernoff bound, one shows that at most a tiny fraction of the sets $S \in \mathcal{S}_n$ are bad for C . (The tiny fraction is $2^{-2^{n/4}}$. Details are in the book.)

Next, one argues that there are at most $2^{2^{n/4}}$ circuits of size $2^{n/8}$. (This is by a counting argument. Details are not in the book and should be verified.) From this, it follows that there is at least one set $S_n \in \mathcal{S}_n$ which is not bad for any such circuit. Fix such a set.

Now, let X_n be uniformly distributed over S_n . Observe that the following three quantities are all the same: the expected value of $C(X_n)$, $\Pr[C(X_n) = 1]$, and $f_C(S_n)$. Since these equalities hold for all circuits C of size at most $2^{n/8}$, it follows that the absolute difference between the behavior of any such circuit on U_n and on X_n is at most $2^{-n/8}$, which grows more slowly than $1/p(n)$ for any polynomial $p(\cdot)$. Hence, the probabilistic ensembles U and X are indistinguishable by polynomial-size circuits. This implies polynomial-time indistinguishability by probabilistic polynomial-time Turing machines.