Mathematical Methods in Computer Science: Exercise 3

Gil Kalai & Avi Wigderson

March 6, 2003

Exercise 1. Here is a model for weak random sources. A distribution X_n on $n - bit$ strings is called a k-source if no string has probability more than 2^{-k} to occur. Note that in particular this means that the Shannon entropy of the source is at least k, but actually our requirement is stronger.

The general problem of using such sources effectively in randomized computations is extremely interesting. In particular, one may ask if there is a deterministic way to extract "nearly" random bits from such sources. Here are negative an positive answers you should prove.

One source - negative Prove that for every function on n bits f_n, there is a k-source X_n with $k = n - 1$ and with $f_n(X_n)$ constant with probability one.

Two independent source - positive existential result Prove that there exist functions f_n on $2n$ bits, such that for every two k-sources X_n, Y_n with $k > 10 \log n$,

$$\left| \Pr[f_n(X_n, Y_n) = 1] - 1/2 \right| < \exp(-\Omega(k))$$

Two independent sources - positive explicit construction Give a polynomial time computable Boolean function f_n on $2n$-bit strings, such that for every two k-sources X_n, Y_n with $k > (\frac{1}{2} + \epsilon)n$ for some constant $\epsilon > 0$,

$$\left| \Pr[f_n(X_n, Y_n) = 1] - 1/2 \right| < \exp(-\Omega(n))$$

Hint: to do the last part, you may want to

1. Prove that every k-source is a convex combination of flat k-sources, namely those who are uniformly distributed over sets of strings of size 2^k exactly.

2. Relate this problem to the discrepancy version of the Bipartite Ramsey problem we solved using Hadamard matrices.

Exercise 2. Let V be a random variable over n-bit strings which is ϵ-biased.

- Prove the best lower bound you can on the size of the support of this distribution.

- Prove that all nontrivial discrete Fourier coefficients of this distribution (in the group Z_2^n) are bounded above by ϵ in absolute value.

- Prove that the L_1 distance of this distribution from the uniform distribution on all n-bit strings is at most $\epsilon 2^n$.