Efficient Generation of Fair Bits

John Gill
Electrical Engineering Department Stanford University

Von Neumann in 1951 described a simple procedure for producing independent, equiprobable ("fair") bits given a source of independent but biased bits. Von Neumann's procedure mapped pairs of input bits to strings of input bits:

$$01 \to 0\,, \quad 10 \to 1\,, \quad 00, 11 \to \Lambda = \text{empty string}$$

The output bits are equiprobable because $\mathrm{P}(01) = \mathrm{P}(10) = pq$, where $p = \mathrm{P}(1)$; they are independent because they are functions of different independent inputs.

Von Neumann's procedure is inefficient; it produces on average pq output bits per input. For example, when $p = 1/2$ the output rate is only $1/4$. Elias and Gill independently in 1972 showed how to increase the efficiency by mapping blocks of n input bits to variable-length strings of output bits. This procedure essentially attains the fundamental upper bound on efficiency—the entropy of the input bits $H(p) = p \log_2(1/p) + q \log_2(1/q)$—for large n.

Both Elias and Gill stated without proof that the generalization of their procedure to inputs obtained from arbitrary finite state Markov processes was trivial; both were wrong. Manuel Blum in 1986 demonstrated that although the output bits were unbiased, they were not always independent. A surprisingly simple modification to the obvious generalized procedure guarantees independence. In this presentation, I will describe Blum's algorithm. Time permitting, I will discuss the reverse problem and describe Knuth and Yao's optimal procedure for generating biased bits from fair bits.