Estimating Parameters of Monotone Boolean Functions
(abstract)

Michael J. Fischer

Department of Computer Science, P.O. Box 208285,
Yale University, New Haven CT 06520-8285, USA
fischer-michael@cs.yale.edu
http://www.cs.yale.edu/~fischer/

Let $F : \{0,1\}^n \to \{0,1\}$ be a monotone Boolean function. For a vector $\boldsymbol{x} \in \{0,1\}^n, let\#(\boldsymbol{x})$ be the number of 1's in $\boldsymbol{x}$, and let $S_k = \{\boldsymbol{x} \in \{0,1\}^n \mid \#(\boldsymbol{x}) = k\}$. For a multiset set $R \subseteq \{0,1\}^n$,define the F-density of R to be

$$D(R) = \frac{\mid \{\boldsymbol{x} \in R \mid F(\boldsymbol{x}) = 1\} \mid}{\mid R \mid}$$

Thus, $D(R) = \mathsf{prob}[F(X) = 1]$, where X is uniformly distributed over R.

Let R_k be a random multiset consisting of m independent samples drawn uniformly from S_k. Then the random variable $Y_k = D(R_k)$ is a sufficient estimator for $D(S_k)$. A naive algorithm to compute $Y_0, \cdots, Y_n$ evaluates F on each of the $m(n+1)$ random samples in $\bigcup_k R_k$ and then computes each $D(R_k)$.

The following theorem shows that the number of evaluations of F can be greatly reduced.

Main Theorem. *There is a randomized algorithm for computing $Y_0, \cdots, Y_n$ that performs at most $m\lceil \log_2(n+2) \rceil$ evaduations of F.*

When n is large and F takes a considerable amount of time to evaluate, this theorem permits a dramatic decrease in the time to compute the Y_k's and hence to estimate the parameters $D(S_k)$ to a given degree of accuracy.

The problem of estimating the parameters $D(S_k)$ arises in the study of error-correcting codes. The vector $\boldsymbol{x} \in \{0,1\}^n$ represents a particular error pattern, and $F(x) = 1$ iff the error-correction algorithm fails to correct all errors. Many error correction algorithms are monotone in the sense that additional errors can never turn an uncorrectable error pattern into a correctable one. In this context, $D(S_k)$ is the probability that a corrupted code word containing k independent bit errors is uncorrectable by the algorithm. Regarded as a function of k, it is a useful measure of the overall error-correction ability of an algorithm.

Wen-Lian Hsu, Ming-Yang Kao (Eds.): COCOON'98, LNCS 1449, p. 3, 1998.