Optimal Separation and Strong Direct Sum for Randomized Query Complexity

Eric Blais, Joshua Brody

Model

- Function $f : \mathcal{X}^n \to \{0,1\}$, $\mathcal{X}$ is finite.
- Error parameter $\varepsilon \geq 0$.
- Query cost $|\mathcal{A}|$ of an algorithm $\mathcal{A}$ is the maximum of coordinates of x queried by $\mathcal{A}$.

Randomized Complexity

- Randomized algorithm $\mathcal{A}$ computes f with an error ε if for every $x \in \mathcal{X}$ holds that
$$\Pr_r[\mathcal{A}(x) = f(x)] \geq 1 - \varepsilon.$$
- Query complexity $R_\varepsilon(f)$ of f is a query cost of the optimal algorithm which computes f with an error ε; $R(f) = R_{1/3}(f)$.
- Average query complexity $\overline{R}_\varepsilon(f)$ of f is an average query cost of the optimal algorithm which computes f with an error ε; $\overline{R}(f) = \overline{R}_{1/3}(f)$.

Distributional Complexity

- Distribution of input μ.
- Deterministic algorithm $\mathcal{D}$ computes f with an error ε if
$$\Pr_\mu[\mathcal{D}(x) = f(x)] \geq 1 - \varepsilon.$$
- Distributional complexity $D^\mu_\varepsilon(f)$ is a query cost of the optimal deterministic algorithm which computes f with an error ε.

Aborting Algorithm

- Algorithms also can abort with probability δ.
- Measures $\overline{R}_{\delta,\varepsilon}(f), R_{\delta,\varepsilon}(f), D^\mu_{\delta,\varepsilon}(f)$ defined similarly, algorithms which can abort are also considered.

Results

Theorem 1 (Error Separation)**.** *For infinitely many values of n and every $2^{-(\frac{n}{\log n})^{1/3}} < \varepsilon \leq \frac{1}{3}$, there exists a total function $f : \{0,1\}^n \to \{0,1\}$ such that*
$$\overline{R}_\varepsilon(f) \geq \Omega\Big(R(f) \cdot \log \frac{1}{\varepsilon}\Big).$$
Theorem 2 (Direct Sum)**.** *For every function $f : \{0,1\}^n \to \{0,1\}$, every $k \geq 2$ and every $0 \leq \varepsilon \leq \frac{1}{20}$ holds that*
$$\overline{R}_\varepsilon(f^k) \geq \Omega\big(k \cdot \overline{R}_{\frac{\varepsilon}{k}}(f)\big).$$

Direct Sum

Lemma 3. *For every function $f : \{0,1\}^n \to \{0,1\}$, every $0 \le \varepsilon < \frac{1}{2}$ and every $0 < \delta < 1$ holds that*

$$\delta \cdot R_{\delta,\varepsilon}(f) \le \overline{R}_\varepsilon(f) \le \frac{1}{1-\delta} \cdot R_{\delta,(1-\delta)\varepsilon}(f).$$

Lemma 4. *For every function $f : \{0,1\}^n \to \{0,1\}$ and any $\alpha, \beta > 0$ such that $\alpha + \beta \le 1$ holds that*

$$\max_\mu D^\mu_{\frac{\delta}{\alpha},\frac{\varepsilon}{\beta}}(f) \le R_{\delta,\varepsilon} \le \max_\mu D^\mu_{\alpha\delta,\beta\varepsilon}(f)$$

Lemma 5. *For every function $f : \{0,1\}^n \to \{0,1\}$, every distribution μ on $\{0,1\}^n$ and every $0 \le \delta, \varepsilon \le \frac{1}{4}$ holds that*

$$D^\mu_{\delta,\varepsilon}(f^k) \ge \Omega\big(k \cdot D^\mu_{\frac{1}{10}+4\delta+4\varepsilon,\frac{48\varepsilon}{k}}(f)\big).$$

Error Separation

Definition 6 (Joining Function)**.** Let $f : \{0,1\}^n \to \{0,1\}$ and $g : \{0,1\}^m \to \{0,1\}$. We define a function $f \circ g : \{0,1\}^{n\times m} \to \{0,1\}$ as

$$f \circ g(x_1, \ldots, x_n) = f\big(g(x_1), \ldots, g(x_n)\big),$$

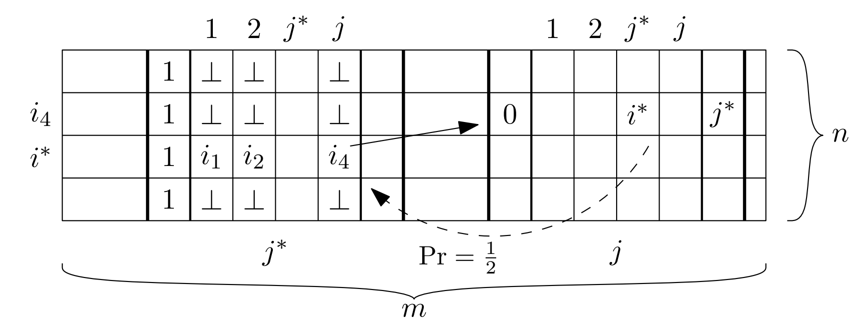
where $x_i \in \{0,1\}^m$.

Definition 7 (Resilient Function)**.** A function $\phi : \{0,1\}^n \to \{0,1\}^m$ is t-resilient for some $1 \le t \le n$ if for any set $S \subseteq [n]$ of $|S| \le t$ of coordinates and any assignment of values for the inputs $\{x_i\}_{i\in S}$, when the values $\{x_i\}_{i\notin S}$ are set uniformly at random then $\phi(x)$ is uniformly distributed in $\{0,1\}^m$.

Theorem 8 (Chor et al.)**.** *For every large enough n, there is a function $\phi : \{0,1\}^n \to \{0,1\}^m$ that is $\frac{n}{3}$-resilient and satisfies $m \ge 0.08n$.*

Functions

- $\textsc{Ptr} : \Gamma^{n\times m} \to \{0,1\}, \Gamma = \{0,1\} \times \big([n] \cup \{\perp\}\big)^m \times \big([m] \cup \{\perp\}\big)$.



- $\textsc{GapZ} : \{0,1\}^m \to \{0,1\}$.

$$\textsc{GapZ}(x) = \begin{cases} 1 & |x| = 0 \\ 0 & |x| = \frac{m}{2} \\ \text{undefined} & \text{otherwise} \end{cases}$$

- $\mathrm{BR} : \Sigma^{n\times m} \to \{0,1\}, \Sigma = \{\textsc{Blue}, \textsc{Red}, \textsc{NotColored}\}$, $x \in \Sigma^{n\times m}$ is valid if each column has exactly 1 colored entry.

$$\mathrm{BR}(x) = \begin{cases} 1 & x \text{ is valid and each colored entry is red} \\ 0 & x \text{ is valid and exactly } \frac{m}{2} \text{ colored entries are blue} \\ \text{undefined} & \text{otherwise} \end{cases}$$