Foundations and Frontiers of Probabilistic Proofs (July 2023)

Worksheet 6: Intro to PCPs

Date: July 19, 2023

Problem 1. (Identity testing) Fix an arbitrary string $s \in \Sigma^n$ and $\varepsilon = \varepsilon(n) \in (0,1)$. Show that the query complexity of detecting whether an unknown string $x \in \Sigma^n$ is equal to s or differs from s in at least an ε fraction of locations is $\Theta(1/\varepsilon)$. That is:

1. Construct an algorithm that makes $O(1/\varepsilon)$ queries to x, and always accepts if $x = s$ and rejects with probability at least $2/3$ if x is ε-far from s.
2. Argue that no algorithm making $o(1/\varepsilon)$ queries satisfies both conditions.

Recall that a language $\mathcal{L}$ has a PCP with perfect completeness, alphabet Σ, proof length l and soundness error ϵ if there exists an algorithm V (the *verifier*) satisfying the following conditions.

- **Completeness:** If $x \in \mathcal{L}$, then there exists $\pi \in \Sigma^{\mathsf{l}}$ such that $\Pr[V^{\pi}(x) = 1] = 1$;
- **Soundness:** If $x \notin \mathcal{L}$, then for every $\tilde{\pi} \in \Sigma^{\mathsf{l}}$, $\Pr[V^{\tilde{\pi}}(x) = 1] \leq \epsilon$.

The query complexity q is the maximum number of queries made by V to its proof string, while the randomness complexity r is the number of coin tosses performed by the verifier.

Below $\mathcal{L}$ is a language that has a PCP with the aforementioned parameters.

Problem 2. (From many to 2 queries) Prove that $\mathcal{L}$ has a PCP with perfect completeness, soundness error $1 - \frac{1-\epsilon}{\mathsf{q}}$, alphabet Σ^{q}, proof length $\mathsf{l} + 2^{\mathsf{r}}$, and query complexity 2. (In other words, one can always reduce query complexity to 2, incurring a loss in soundness error and alphabet size.)

Problem 3. (Lower bound on soundness error) Suppose that there exists $\mathbb{x} \notin \mathcal{L}$ such that for every choice of verifier randomness $\rho \in \{0,1\}^{\mathsf{r}}$ there exists a proof $\pi \in \Sigma^{\mathsf{l}}$ such that $V^{\pi}(\mathbb{x}; \rho) = 1$. Prove that $\epsilon \geq 2^{-\mathsf{q} \log |\Sigma|}$.

Problem 4. (More on lower bounds) The *Exponential Time Hypothesis* (ETH) states that 3SAT cannot be decided by any deterministic algorithm running in time $2^{o(n)}$. Prove that, assuming ETH, if $\mathcal{L} = \text{3SAT}$ and $\mathsf{r} + \mathsf{q} \log |\Sigma| = o(n)$, then $\epsilon \geq 2^{-\mathsf{q} \log |\Sigma|}$. (*Hint: prove that ETH implies the assumption to the prior problem.*)