6.841: Advanced Complexity Theory Fall 2012

Problem Set 1

Prof. Dana Moshkovitz/TA: Henry Yuen *Due Date: September 20,2012*

Turn in your solution to each problem on a separate piece of paper. Mark the top of each sheet with the following: (1) your name, (2) the question number, (3) the names of any people you worked with on the problem, or "Collaborators: none" if you solved the problem individually. We encourage you to spend time on each problem individually before collaborating!

1 Problem 1 – Circuits and the polynomial hierarchy

(a) Prove that $\mathsf{NTIME}(n) \subseteq \mathsf{DTIME}(n^c)$ implies that $\Sigma_2^{\mathsf{P}}\mathsf{TIME}(n) \subseteq \mathsf{NTIME}(n^c)$.

(b) Show that for every k, there exists a language in Σ_2^{P} that does not have circuits of size n^k. [Note: this does not show that PH does not have polynomial sized circuits! Indeed, showing that $\mathsf{PH} \not\subseteq \mathsf{P/poly}$ (or $\mathsf{PSPACE} \not\subseteq \mathsf{P/poly}$, or even $\mathsf{NEXP} \not\subseteq \mathsf{P/poly}$) seems to be quite beyond the reach of current circuit lower bound techniques.]

(c) Here, we will show that upper bounds can sometimes be used to show lower bounds. Suppose that $\mathsf{P} = \mathsf{NP}$[1]. First, show that $\mathsf{P} = \mathsf{NP}$ implies that $\mathsf{EXP} = \mathsf{NEXP}$, where NEXP is the exponential-time version of NP (i.e. the proof size can be $2^{O(n^c)}$ for some constant c, and the proof verifier can also run in exponential time). Then, consider an exponential-time version of the polynomial hierarchy to deduce our lower bound: there exists a language in EXP that requires circuits of size $2^n/n$.

2 Problem 2 – A dramatic collapse!

Show that $\mathsf{PSPACE} \subseteq \mathsf{P/poly}$ implies that $\mathsf{PSPACE} = \Sigma_2^{\mathsf{P}}$. (In other words, simulating small space with small circuits means that polynomial space collapses into a complexity theoretic blackhole.)

3 Problem 3 – NP-completeness of 3SAT does not relativize

Demonstrate an oracle A and a language $L \in \mathsf{NP}^A$ such that L is not polynomial-time reducible to 3SAT, even when the reduction algorithm is given oracle access to A. This shows that the Cook-Levin theorem is nonrelativizing! *Bonus*: Identify the non-relativizing components in the proof of the Cook-Levin theorem.

[1] Some believe this not to be true.

4 Problem 4 – Intractability of computing Vapnik-Chervonenkis dimension

Optional problem. An important concept in machine learning and computational learning theory is that of Vapnik-Chervonenkis (VC) dimension: Let $\mathcal{S} = \{S_1, S_2, \ldots, S_m\}$ be a family of subsets in a finite universe $\mathcal{U}$. The VC dimension of $\mathcal{S}$, denoted $VC(\mathcal{S})$, is the size of the largest set $X \subseteq \mathcal{U}$ such that for every $X' \subseteq X$, there exists an i for which $S_i \cap X = X'$ (we say that X is *shattered* by $\mathcal{S}$).

Often, the set $\mathcal{S}$ will represent a bunch of hypotheses; X will represent some training data. If X is shattered by $\mathcal{S}$, that means that all possible binary classifications of points in X have a consistent hypothesis in $\mathcal{S}$. If X is very large, then this indicates that the set of hypotheses $\mathcal{S}$ is very *expressive*; it contains very complex explanations. Thus, intuitively, the VC-dimension of a set of hypotheses is a measure of its descriptive complexity.

Suppose we had a family of sets $\mathcal{S} = \{S_1, \ldots, S_m\}$ that was implicitly represented by a circuit: for all i, $x \in S_i$ if and only if $C(i, x) = 1$. The following language captures the problem of computing VC dimension:

$$\text{VC-DIMENSION} = \{\langle C, k\rangle \mid C \text{ represents a set family } \mathcal{S} \text{ s.t. } VC(\mathcal{S}) \geq k\ \}.$$

Note that k is written in *binary*. Show that VC-DIMENSION is Σ_3^{P}-complete. [Hint: You can use the fact that Σ_3^{P}-SAT (the generalization of CNF-SAT to include $\exists \ldots \forall \ldots \exists$ quantifiers) is complete for Σ_3^{P}.]