Final

due Friday March 18, 2022

Directions: Problems 1−4 are each 10 points; problem 5 is 15 points.

Part 1 (QEM)

Problem 1. Suppose $f : \{0,1\}^* \to \{0,1\}$ can be computed by a 2−round interactive proof system $\langle \mathsf{P}, \mathsf{V} \rangle$ with perfect completeness. Prove that f can be computed by a 3−round public-coin interactive proof system $\langle \mathsf{P}', \mathsf{V}' \rangle$ perfect completeness.

Problem 2. In class we proved that $\mathsf{BPP} \subset \Sigma_2$ (Sipser-Gács theorem) so $\mathsf{BPP} \subset \mathsf{PSPACE}$ (since $\mathsf{PH} \subset \mathsf{PSPACE}$). Prove directly that $\mathsf{BPP} \subset \mathsf{PSPACE}$ by showing that any $f \in \mathsf{BPP}$ can be computed by a polynomial space TM (no credit will be given for invoking Sipser-Gács).

Part 2

The problems in this section will walk you through the proof that the "large graph partition" problem is NP−hard. Define the function $f_{\mathsf{LGP}} : \{0,1\}^* \to \{0,1\}$ as follows: on input (G, M) where $G = (V, E)$ is a graph and $M \in \mathbb{N}$ is a positive integer, $f_{\mathsf{LGP}}(G, M)$ outputs 1 if there exists $S \subset V$ such that

$$\Big\{ e = (v, w) \in E : v \in S \ \&\ w \notin S \Big\} \geq M,$$

and outputs 0 if not. The proof that f_{LGP} is NP−hard will be via a reduction from $\mathsf{3SAT}$, and will go through the intermediate "not-unanymous" problem. The input to the k−not-unanymous problem is a formula Φ over variables $\{x_1, \ldots, x_n\}$ and with clauses $\varphi_1, \ldots, \varphi_m$ where each clause contains k literals. An example of a clause when $k = 3$ is $(x_i, \overline{x}_j, x_k)$. The function $f_{\mathsf{kNU}} : \{0,1\}^* \to \{0,1\}$ takes input Φ and outputs 1 if there exists an assignment to the variables such that the literals in every clause are not all equal. The next two problems show that f_{LGP} is NP−hard, except for one missing step; namely that $f_{\mathsf{4NU}} \leq f_{\mathsf{3NU}}$. This can be proved in precisely the same way as $f_{\mathsf{4SAT}} \leq f_{\mathsf{3SAT}}$ (and you do not have to prove this).

Problem 3. Prove that $f_{\mathsf{3SAT}} \leq f_{\mathsf{4NU}}$.

Problem 4. Prove that $f_{\mathsf{3NU}} \leq f_{\mathsf{LGP}}$.

Part 3

In class we saw the FGLSS construction to convert a 2-round 2-prover MIP system into a graph. The exercises in this section will work through the analysis of (a simple case of) this construction. Suppose $f : \{0,1\}^* \to \{0,1\}$ is a function which is computed by a $2-$round, $2-$prover MIP system $\langle \mathsf{P}_1, \mathsf{P}_2, \mathsf{V}\rangle$ with perfect completeness and soundness $s < 1$. So specifically, for all $\mathbf{x} \in \{0,1\}^*$ such that $f(\mathbf{x}) = 0$, $\Pr\big[\langle \mathsf{P}_1, \mathsf{P}_2, \mathsf{V}\rangle(\mathbf{x}) = 1\big] \le s$. Denote the transcript of the MIP system as $\big((a_1, b_1), (a_2, b_2)\big)$ where in the first round V sends a_i to P_i and in the second round, receives b_i from P_i; then outputs the bit $\mathsf{V}_{\mathsf{out}}\big(\mathbf{x}, (a_1, b_1), (a_2, b_2)\big)$. We assume V's message in the first round is computed as the output of a probabilistic computation where V first chooses a random string $r \sim \{0,1\}^m$ and then outputs $(a_1, a_2) = \mathsf{V}_1(r)$. We assume $b_1, b_2 \in \{0,1\}^\ell$.

Now, given $\mathbf{x} \in \{0,1\}^*$, the FGLSS construction outputs the graph $G_{\mathbf{x}}$ where the vertices and edges are defined as follows:

- **Vertices:** There is one vertex in $G_{\mathbf{x}}$ for each "accepting transcript", *i.e.*, for each $\big((a_1, b_1), (a_2, b_2)\big)$ such that $\mathsf{V}_{\mathsf{out}}\big(\mathbf{x}, (a_1, b_1), (a_2, b_2)\big) = 1$. We denote the vertex corresponding to $\big((a_1, b_1), (a_2, b_2)\big)$ as v_{a_1,b_1,a_2,b_2}.
- **Edges:** Vertices are connected as long as their underlying transcripts are not inconsistent.
  - The transcripts $\big((a_1, b_1), (a_2, b_2)\big)$ and $\big((a_1', b_1'), (a_2', b_2')\big)$ are inconsistent if $a_i = a_i'$ and $b_i \ne b_i'$ for some $i = 1, 2$.

Problem 5. This problem has 3 parts, each worth 5 points.

(a) Prove that if $f(\mathbf{x}) = 1$ then $G_{\mathbf{x}}$ has a clique of size 2^m.

(b) Prove that if $G_{\mathbf{x}}$ has a clique of size $\sigma \cdot 2^m$, then there exist cheating prover strategies P_1^* and P_2^* such that $\Pr\big[\langle \mathsf{P}_1^*, \mathsf{P}_2^*, \mathsf{V}\rangle(\mathbf{x}) = 1\big] = \sigma$.

(c) Parts (a) and (b) say, respectively, that if $f(\mathbf{x}) = 1$ then $G_{\mathbf{x}}$ has a clique of size 2^m, while if $f(\mathbf{x}) = 0$ then the largest clique in $G_{\mathbf{x}}$ has size $s \cdot 2^m$. Deduce that if an $\mathsf{NP}-$hard function f can be computed by a $2-$round, $2-$prover MIP system with perfect completeness and soundness s, then it is also $\mathsf{NP}-$hard to distinguiush, given a graph G and a bound B, whether G has a clique of size B, or whether the largest clique in G has size at most sB.