CS 2429
Propositional Proof Complexity
ASSIGNMENT # 1
Due: April 6, 2017

1. Prove that NP equals coNP if and only if there exists a polynomially bounded propositional proof system. You can assume that SAT is NP-complete, but do not assume that UNSAT is coNP-complete.

2. Recall that a CNF formula is Horn if each clause contains at most one positive literal. Prove that any unsatisfiable Horn CNF formula has a polynomial-size Resolution refutation. Hint: Notice that if every clause has size at least two, then the formula is satisfiable by the all zero assignment. Construct a specific DPLL tree for an unsatisfiable Horn formula using this hint, and prove that the size of the DPLL tree produced is linear in the number of underlying variables.

3. The Tseitin tautologies, TS_n on the complete graph with n vertices are as follows. The underlying variables are $e_{i,j}$ for all $i, j \leq n$, $i \neq j$. Note that the variable $e_{i,j}$ has *unordered* subscripts– that is, it is the same variable as $e_{j,i}$. Corresponding to each vertex i is a set of clauses that states that the sum of the variables $e_{i,j}$, $j \leq n$, is odd. Note that whenever n is odd, this set of clauses is unsatisfiable since we have n equations, each expressing that the mod 2 sum of a group of variables is odd, and such that each variable occurs exactly twice in the n equations. We express TS_n in CNF form as follows. For each $i \leq n$, there are 2^{n-2} clauses of the form $\vee_{j \neq i} e_{i,j}^{b_j}$ where $e_{i,j}^0 = e_{i,j}$ and $e_{i,j}^1 = \neg e_{i,j}$ and $b_1 + \ldots + b_n = 0 \ mod\ 2$. Note that the number of variables is $O(n^2)$ and the size of TS_n is $O(n2^n)$.

   For example, for $n = 3$, we have variables $e_{1,2}, e_{1,3}, e_{2,3}$ and the clauses are as follows: $(e_{1,2} \vee e_{1,3})(\neg e_{1,2} \vee \neg e_{1,3})(e_{2,3} \vee e_{2,1})(\neg e_{2,3} \vee \neg e_{2,1})(e_{1,3} \vee e_{1,2})(\neg e_{1,3} \vee \neg e_{1,2})$

   - Call a clause ϵ-*wide* if it contains ϵn^2 variables. Prove that there exists an ϵ such that for sufficiently large odd n, *any* Resolution refutation of TS_n has an ϵ-wide clause.
   - Prove using the your wide clause lemma, and the size-width theorem for Resolution, a lower bound of the form $2^{\delta n^2}$, which is quasipolynomial in the size of TS_n.

4. This question is about converting a "game" characterization of a Frege proof to a sequent calculus proof.

   Let $f = C_1 \wedge \ldots \wedge C_m$ be an unsatisfiable CNF over $x_1, \ldots, x_n$ Suppose that f has a Frege proof as characterized by the liar/prover game. I.e., there is a decision tree T where each node queries a formula over $x_1, \ldots, x_n$, and at the leaves we have truth table contradictions.

Show how to convert T to a PK (sequent calculus) proof of the sequent $C_1, \ldots, C_m \rightarrow$.

Hint. Label the root of the tree T with $C_1, \ldots, C_m \rightarrow$ and think of each node in the tree as an application of the cut rule. So if the first node queries some formula f, then we would label the left node ($f = 1$) with $f, C_1, \ldots, C_m \rightarrow$ and the right node ($f = 0$) with $C_1, \ldots, C_m \rightarrow f$ Thus if a path to a leaf has queries $f_1 = 1, f_2 = 1, ..f_k = 1, g_1 = 0, ..g_k = 0$, then that leaf would be labelled by the sequent $f_1, .., f_k, C_1, ..., C_m \rightarrow g_1, \ldots, g_k$.

For each such leaf, show how to derive this line from the axioms and rules (using the fact that we have a truth table contradiction).