Homework 5
Due Date: March 28, 2019 at 11:59 PM on Gradescope.

Question 1
Consider the following two decision problems:

- **OR(G,k):** Either $G$ contains no odd length cycle OR $G$ has a vertex cover $S \subseteq V(G)$ of size $|S| \leq k$.

- **AND(G,k):** $G$ contains no odd length cycle AND $G$ has a vertex cover $S \subseteq V(G)$ of size $|S| \leq k$.

One of the problems is NP-Complete and the other one is in P. Your mission – you better choose to accept it – is to figure out which of the above problems is NP-Complete. You should prove that your answer is correct by providing a polynomial time algorithm to solve one problem and by proving that the other problem is NP-Complete.

Question 2
A sorted sequence of real numbers is a sequence of the form $a_1 \leq a_2 \leq \ldots \leq a_n$. A sequence of real numbers $b_1, b_2, \ldots, b_n$ is said to be $\varepsilon$-close to being sorted if there is a sorted sequence $a_1 \leq a_2 \leq \ldots \leq a_n$ such that $\sum |a_i - b_i| \leq \varepsilon$.

Of course, you know what a sorted sequence of numbers looks like. Unfortunately, your manager at a software development company does not. He has a sequence of real numbers $Y = y_1, y_2, \ldots, y_n$ which is not quite sorted which will be fed as an input to TheGreatExperiment. There is some hope that this will not lead to any disaster as long as the sequence $Y$ is (pretty darn) $\varepsilon$-close for some tiny $\varepsilon > 0$. Use linear programming to determine whether the sequence $Y$ is $\varepsilon$-close to being sorted.

Question 3
Recall that a 2-SAT Instance $\Phi$ consists of $n$ variables $x_1, \ldots, x_n$ and $m$ clauses $C_1, \ldots, C_m$ where each clause $C_i = \ell_{i_1} \lor \ell_{i_2}$ is the disjunction of two distinct literals. In this problem you will develop a linear time algorithm to solve the problem by reducing the 2-SAT instance to a directed graph $G_\Phi$ defined as follows:

**Nodes:** For each literal $\ell$ we add a node $v_\ell$ ($2n$ nodes total)

**Notation:** Given a node $v_\ell$ we use $v_{\overline{\ell}}$ to denote the node corresponding to $\overline{\ell}$. For example
if $\ell = \overline{x}$ then $v_\ell = v_x$.

**Edges:** For each clause $C = \ell \lor \ell'$ we add two directed edges $(v_\ell, v_{\ell'})$ and $(v_{\ell'}, v_\ell)$.

**Part 0** Suppose that we have a directed path $P = v_{\ell_1}, v_{\ell_2}, \ldots, v_{\ell_k}$ in $G_\Phi$. Explain why $P = v_{\ell_k}, \ldots, v_{\ell_2}, v_{\ell_1}$ is a valid directed path in $G_\Phi$.

**Part 1:** Suppose that for some literal $\ell$ there is directed path from node $v_\ell$ to $v_{\ell'}$ in $G_\Phi$ and that there is also a directed path from $v_{\ell'}$ to $v_\ell$ in $G_\Phi$. Show that the 2-SAT instance $\Phi$ is not satisfiable.

**Part 2:** Suppose that for every literal $\ell$ the nodes $v_\ell$ and $v_{\overline{\ell}}$ are in separate strongly connected components in $G_\Phi$. Show that the 2-SAT instance $\Phi$ is satisfiable.

**Part 3:** Describe an $O(n+m)$ time algorithm which solves the 2-SAT problem. If the 2-SAT formula $\Phi$ is not satisfiable then the algorithm should output no-solution otherwise the algorithm should output a satisfying assignment. (Hint: You may use Tarjan’s Algorithm as a blackbox.).

**Question 4**

So, we are back to city Z which is frequented by villainous monsters. The city has started a Hero Association to deal with these troublemakers. There is a set $V$ of $n$ villains and a set $H$ of $m$ heroes. Each hero $h \in H$ is a resource that can be allocated to fight off exactly one villain. Each villain $v \in V$ has a subset $H_v \subseteq H$ of heroes that must be assigned to capture $v$. The Hero Association needs you to solve the decision problem $ZHERO(n, m, k)$ which asks whether they can allocate the heroes to villains such that at least $k$ villains get captured.

Let us try to get a feel for how this problem behaves through the following subparts.

1. Assume $k = O(1)$ and $m$ and $n$ are large. Help the Hero Association come up with a polynomial time algorithm to solve $ZHERO(n, m, k)$ in this setting.

2. Assume all of $n, m, k$ are all very large. Answer the following questions.

   (a) Show that $ZHERO(n, m, k)$ is in $\text{NP}$.

   (b) Show that $ZHERO(n, m, k)$ is $\text{NP}$-Hard by giving a reduction from the matching problem in 3-uniform hypergraphs. The matching problem in 3-uniform hypergraphs is as follows: You are given a triple of vertex sets $A, B, C$ where $|A| = |B| = |C| = m$. You are told that a subset $E \subseteq A \times B \times C$ form hyperedges each of which span 3 vertices – one each from $A$, $B$ and $C$. We want to decide whether there exists a subset $M \subseteq E$ with $|M| \geq k$ where no two hyperedges in $M$ share any vertex.

   (c) Assume you have an oracle which solves the decision problem $ZHERO(n, m, k)$. Give an algorithm that calls this oracle a polynomial number of times and finds the biggest subset $S \subseteq V$ of villains that can be safely captured.