- You may use any of the following known NP-complete problems to show that a given problem is NP-complete: 3-SAT, INDEPENDENT-SET, VERTEX-COVER, SET-COVER, HAMILTONIAN- CYCLE, HAMILTONIAN-PATH, SUBSET-SUM, 3-COLORING.

There are 6 questions for a total of 100 points.

---

1. (*PCP and Hardness of approximation*) In this question, we will use r to denote the number of random bits and q the number of queries in the context of $\mathsf{PCP}_{c,s}(r, q)$. Moreover, c denotes completeness and s denotes soundness.
   (a) (10 points) Recall that the PCP theorem says that $\mathsf{PCP}(O(\log n), 3)$ is the same as NP. Discuss the complexity of $\mathsf{PCP}(O(\log n), 2)$.

2. (*PCP and Hardness of approximation*) Using some slightly advanced machinery, the following PCP theorem variant has been shown.

   Theorem 1: For every $\varepsilon > 0$, $\mathsf{NP} = \mathsf{PCP}_{1-\varepsilon,1/2+\varepsilon}(O(\log n), 3)$. Moreover, the PCP verifier uses $O(\log n)$ random bits to compute three positions of the proof, i, j, k, and a bit b and accepts iff $y[i] + y[j] + y[k] = b \ (mod\ 2)$. Here $y[i]$ denotes the i^{th} bit of the proof y.

   Consider the following optimization problem:

   Given m constraints in n 0/1 variables $x_1, ..., x_n$, find an assignment to the variables that maximises the number of satisfied constraints. Every j^{th} constraint is of the form

   $$x_{j_1} + x_{j_2} + x_{j_3} = b_j \ (mod\ 2).$$

   Answer the following questions:
   (a) (5 points) How hard is the problem when $m < n$?
   (b) (5 points) Design a (1/2)-approximation algorithm for this problem.
   (c) (5 points) Show that there cannot exist an efficient $(\frac{1}{2}+\varepsilon)$-approximation algorithm for this problem unless $\mathsf{P} = \mathsf{NP}$. Use Theorem 1 in your argument.
   (d) (5 points) Use the previous part to argue that there cannot exist an efficient $(\frac{7}{8}+\varepsilon)$-approximation algorithm for the MAX-3-SAT problem unless $\mathsf{P} = \mathsf{NP}$. Recall that in the MAX-3-SAT problem, every clause has exactly three distinct literals and the goal is to maximise the number of satisfied clauses. *Recall, we discussed a $\frac{7}{8}$-approximation algorithm for this problem in the class.*

3. (20 points) (*LP relaxation*) Consider the problem of finding the maximum weight perfect matching in a weighted bipartite graph $G = (L, R, E)$ where $|L| = |R|$ and edge weights w_{ij}'s are positive integers. Here is an LP relaxation for this problem:

   $$\text{Maximise} \quad \sum_{i \in L, j \in R} w_{ij} \cdot x_{ij}$$

   Subject to:

   $$\sum_{j \in R} x_{ij} = 1 \quad \text{for all } i \in L$$

   $$\sum_{i \in L} x_{ij} = 1 \quad \text{for all } j \in R$$

   $$0 \leq x_{ij} \leq 1 \quad \text{for all } i \in L \text{ and } j \in R$$

   Argue that there is an integer optimal solution for the above LP relaxation.

4. (20 points) (*Randomized rounding*) Consider the maximum cut problem on directed graphs. Given a directed graph $G = (V, E)$ with positive edge weights $w_{ij} \geq 0$ for every $(i, j) \in E$, partition the nodes V into sets $(S, \bar{S})$ such that the sum of the weight of edges from S to $\bar{S}$ is maximised. Consider the following LP relaxation for this problem:

$$\text{Maximise} \quad \sum_{(i,j)\in E} w_{ij} \cdot x_{ij}$$

Subject to:

$$x_{ij} \leq y_i \quad \text{for all } (i,j) \in E$$
$$x_{ij} \leq 1 - y_j \quad \text{for all } (i,j) \in E$$
$$0 \leq y_i \leq 1 \quad \text{for all } i \in V$$
$$0 \leq x_{ij} \leq 1 \quad \text{for all } (i,j) \in E$$

Let (y^*, x^*) be an optimal solution to the above relaxed LP for the maxcut problem. Consider the cut created by randomly rounding, putting node i into S with probability $(\frac{1}{4} + \frac{y_i^*}{2})$. Show that the expected weight of the cut so constructed is at least half the optimal cut weight.

5. (*LP duality*) A zero-sum game between two players is defined using an $m \times n$ matrix A with a "row" player R and a "column" player C. Every row denotes a strategy of the row player R and every column denotes a strategy for the column player C. For the row player playing $1 \leq i \leq m$ and the column player playing $1 \leq j \leq n$, the "payoff" to the row player is $A[i, j]$ (*i.e., if $A[i, j]$ is positive C pays $A[i, j]$ to R, otherwise, R pays $-A[i, j]$ to C*). The players can play a "mixed" strategy, instead of a "pure" one (i.e., picking a row/column), where they pick a probability distribution over the rows/columns and pick a row/column based on this probability distribution. For example, R can pick $\mathbf{x} \in \mathbb{R}^n$ with $\sum x_i = 1$ and C can pick $\mathbf{y} \in \mathbb{R}^n$ with $\sum y_i = 1$. In this case, the payoff to R from this mixed strategy is $\mathbf{x}^T A\mathbf{y}$. We can make the following observations about mixed strategies:

1. The best mixed strategy for R is given by: $\max_{\mathbf{x}} \min_{\mathbf{y}} \mathbf{x}^T A\mathbf{y}$
2. The best mixed strategy for C is given by: $\min_{\mathbf{y}} \max_{\mathbf{x}} \mathbf{x}^T A\mathbf{y}$

You will be asked to prove the following claim.

<u>*Claim 1*</u>: Show that for any fixed mixed strategy $\mathbf{x}$ for R, $\min_{\mathbf{y}} \mathbf{x}^T A\mathbf{y}$ is attained for a pure strategy of C. Similarly, for any fixed mixed strategy $\mathbf{y}$ for C, $\max_{\mathbf{x}} \mathbf{x}^T A\mathbf{y}$ is attained for a pure strategy of R.

Using the above claim, we get that:

1. The best mixed strategy for R is given by: $\max_{\mathbf{x}} \min_j \sum_{i=1}^{m} A[i, j]x_i$.
2. The best mixed strategy for C is given by: $\min_{\mathbf{y}} \max_i \sum_{j=1}^{n} A[i, j]y_j$.

We note that R's best mixed strategy can be found by solving the following LP:

$$\text{Maximise} \quad z$$

Subject to:

$$z - \sum_{i=1}^{m} A[i,j]x_i \geq 0, \quad \text{for } j = 1, ..., n$$
$$\sum_{i=1}^{m} x_i = 1$$
$$x_i \geq 0 \quad \text{for } i = 1, ..., m$$

You need to do the following for this question:

(a) (5 points) Prove claim 1.

(b) (5 points) Show that the dual of the above LP computes the best mixed strategy for C.

Using the duality theorem, we can now conclude that

$$\max_{\mathbf{x}} \min_{\mathbf{y}} \mathbf{x}^T A \mathbf{y} = \min_{\mathbf{y}} \max_{\mathbf{x}} \mathbf{x}^T A \mathbf{y}$$

This is called the *von Neumann's minimax theorem for zero-sum games.*

6. (*Primal-dual*) Consider the following problem defined on sets:

    Given the set of elements $U = \{1, ..., n\}$ with associated non-negative weights $w_1, ..., w_n$. Also given are subsets $T_1, ..., T_m$ of U, each of size at most γ. The goal is to find a subset $S \subseteq U$ of elements with minimum total weight such that for every $1 \leq j \leq m, |S \cap T_j| \geq 1$ (*i.e., there is at least one element from every* T_j *in* S).

    (a) (5 points) Show that the problem is NP-hard for $\gamma \geq 2$.

    (b) (15 points) Design an primal-dual based γ-approximation algorithm for this problem. Use ideas similar to those developed in class for the set cover problem. Discuss correctness and running time.