CS 6515/4540 (Fall 2025)

Exam 2

Name: ______________________________

GTID: ______________________________

GT Username: ______________________________

INSTRUCTIONS

1. When the exam starts, write your name (or GT ID or GT Username) on each page.
2. Bringing two pages of notes is permitted.
3. No books, calculators, or other electronic devices permitted.
4. Make your answers clear and concise.
5. ONLY write on the **FRONT** of each sheet. Backs will not be scanned.
6. You may refer to standard algorithms from class without giving full pseudocode.

I am aware of and in accordance with the Academic Honor Code of Georgia Tech and the Georgia Tech Code of Conduct. I will use no person's help on this test. Also, I have read all the instructions on this page.

Signature: ______________________________

1 Short Answer Questions (36 points, each part 6 points)

Unless explicitly mentioned, a short 1–2 line justification is required.

1. **LP.** Every convex set can be written as an intersection of a finite number of half-spaces (linear inequalities).

   ○ True

   ○ False

   > No. For example, a sphere in $\mathbb{R}^2$ requires infinitely many supporting half-spaces to describe exactly.

2. **Facility Location.**

   Consider the LP relaxation for the facility location problem from class, where I is the set of potential facility locations and J is the set of clients such that $I \cap J = \emptyset$, i.e., no client share a location with a potential facility:

   $$\min \sum_{i \in I} f_i y_i + \sum_{i \in I} \sum_{j \in J} d_{ij} x_{ij}, \quad \text{s.t.} \ \sum_{i \in I} x_{ij} = 1 \ \forall j \in J, \quad x_{ij} \leq y_i \ \forall i, j, \quad x_{ij}, y_i \geq 0.$$

   Suppose all opening costs are zero, i.e., $f_i = 0$ for all i. What happens to the optimal LP value?

   (A) The optimal LP value is 0.

   (B) Every client connects to its nearest facility.

   (C) The LP becomes unbounded.

   (D) The LP value equals the total connection cost of the MST.

   > **Answer: (B).** When opening facilities is free, the optimal solution opens every facility needed to minimize connection costs. Each client connects to the nearest open facility.

3. **Random Vars.** For two independent random variables X, Y, do we always have $\mathbb{E}[2^{X+Y}] = \mathbb{E}[2^X] \cdot \mathbb{E}[2^Y]$? Briefly explain why, or give a counterexample.

   > True. 2^X and 2^Y are independent and the expectation of product of two independent r.v. is the same as the product of the expectations.

4. **Farkas' Lemma.** Let $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, $b = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$. Use Farkas to show $Ax = b, x \geq 0$ is infeasible by giving a suitable λ_1, λ_2.

   Take $\lambda = (1, 1)$. Then $\lambda^\top A = (0, 0) \geq 0$ and $\lambda^\top b = -2 < 0$, which certifies infeasibility.

5. **Probability.** Suppose you are writing an algorithm that first calls function A and then calls function B. Both functions A and B have a randomized implementation but your algorithm does not use any other randomness.

   If functions A and B are independent of each other (i.e., they don't share randomness), which of the following is most accurate about the failure probability of your algorithm?

   (A) Pr[Failure of A] + Pr[Failure of B].

   (B) Pr[Failure of A] $\times$ Pr[Failure of B].

   (C) 1- $\Big($1-Pr[Failure of A]$\Big)$ $\times$ $\Big($1-Pr[Failure of B]$\Big)$.

   (D) None of the above

   C. Since A and B are independent functions, we know the probability that both of them succeed is the product of the probabilities that each succeed. So $\Pr(\text{success}) = (1 - \Pr(\text{Failure of } A)) \times (1 - \Pr(\text{Failure of } B))$.

6. **Set Cover.** For the set cover problem, suppose after solving the LP relaxation we get a fractional solution with $x_k^* = 1/3$ for every set $S_k \in \mathcal{S}$. Suppose we perform randomized rounding where we take each set S_k independently with probability 1/3 (i.e., without scaling by factor $10 \log n$). Is any element e guaranteed to be covered with probability at least 1/2?

   True. We know for each element e, $\sum_{e \in S_i} x_i \geq 1$, and since all $x_i = 1/3$, we know that e is contained in at least three sets. So the probability of e not being covered is less than $(1 - 1/3)^3 = 8/27 < 1/2$, and so e is covered w.p. $\geq 1/2$

2 Linear Programming (25 points)

1. **Polygon (10 points).** Let P be a convex polygon given by

$$a_i x + b_i y \le c_i \quad (i = 1, \ldots, m).$$

Given a fixed point $q = (x_0, y_0)$, find $p = (x_1, y_1) \in P$ minimizing

$$\|p - q\|_\infty = \max\{|x_1 - x_0|,\ |y_1 - y_0|\}.$$

Formulate a Linear Program that computes the minimum L_∞ distance from q to P.
(*Hint:* Introduce t with $|x_1 - x_0| \le t$, $|y_1 - y_0| \le t$ and linearize.)

$$\min t$$

$$t \ge x_1 - x_0$$

$$t \ge x_0 - x_1$$

$$t \ge y_1 - y_0$$

$$t \ge y_0 - y_1$$

$$a_i x_1 + b_i y_1 \le c_i \quad (i = 1, \ldots, m).$$

2. **Matching.** Consider the (non-bipartite) maximum matching LP relaxation on $G = (V, E)$ with $\delta(v)$ denoting all the edges incident to any vertex $v \in V$:

$$\max \sum_{e \in E} x_e \quad \text{s.t.} \quad \sum_{e \in \delta(v)} x_e \leq 1 \ \forall v \in V, \qquad x_e \geq 0 \ \forall e \in E.$$

(a) LP Relaxation (5 points) Prove that for any graph $G = (V, E)$, the optimal value of this LP relaxation is at least the size of the maximum matching in G.

> We let $x_e = 1$ if e is in the matching and 0 otherwise. This is feasible: clearly the degree of every vertex is ≤ 1 in a matching, thus $\sum_{e \in \delta(v)} x_e = \deg(v) \leq 1$. Hence all constraints are satisfied and we know the objective function $\sum_e x_e$ is the size of the matching.

(b) Integrality gap (5 points) Give an example of a graph G where the LP relaxation value is strictly more than the size of the maximum matching on G.

(*Hint:* There is a graph already on just 3 vertices.)

> Take a triangle on 3 vertices. Integral optimum picks 1 edge. LP can set $x_e = \frac{1}{2}$ on all three edges in triangle, yielding $\frac{3}{2}$.

(c) Dual of the LP (5 points). Write the dual of the above max-matching LP relaxation.

$$\min \sum_{v \in V} y_v \quad \text{s.t.} \quad y_u + y_v \geq 1 \ \forall \{u, v\} \in E, \qquad y_v \geq 0 \ \forall v \in V.$$

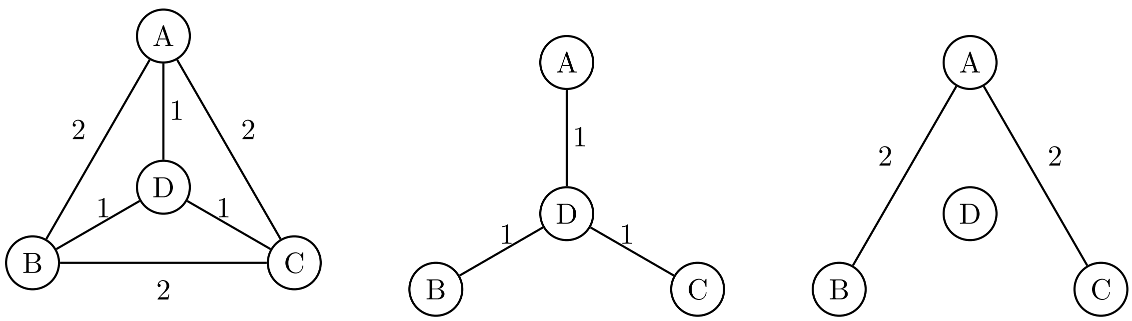
This is the fractional vertex cover LP.

3 Steiner Tree (19 Points)

Suppose we are given a fully-connected (complete) graph $G = (V, E)$ with nonnegative edge weights satisfying the triangle inequality, and a subset $T \subseteq V$ of vertices that must be connected. Find a minimum-weight subset of edges $F \subseteq E$ that connects all vertices in T (vertices in $V \setminus T$ may or may not be connected).
(*Note:* The optimal solution is a tree because otherwise we can drop a cycle edge while maintaining connectivity. Finding the minimum such tree (known as Steiner tree) is NP-complete.)

Let G_T be a fully-connected subgraph of G induced by T (i.e., the subgraph where we remove vertices $V \setminus T$). In this problem we will prove that the minimum spanning tree (MST) of G_T is a 2-approximation to the minimum

Example.



Graph with $T = \{A, B, C\}$ — *Steiner Tree* — MST of G_T

1. (12 points) Let F^* denote the subset of edges that form the minimum Steiner tree. Prove how we can use it to form a spanning tree of G_T (note this spanning tree cannot visit any vertex in $V \setminus T$ since G_T only has vertices T) of total weight at most twice the total weight of F^*.

   (*Hint:* The proof is similar to why MST is a 2-approx for TSP. Recall that G is complete and edge-weights satisfy triangle inequality.)

   We first construct a tour/walk that visits every vertex in T (but may visit other vertices as well) of total weight at most twice the total weight of F^*. This can be achieved by doubling every edge of tree F^* and traversing these edges in DFS (as we did for TSP).

   Now, we can make this tour/walk into a path on vertices of T by short-circuiting: since the walk is a sequence of vertices in V and we only care about vertices in T then if we are on a vertex $u \in T$ then we can just use the direct edge from u to the next distinct vertex $v \in T$ visited by this walk; and by triangle inequality this edge will have lower cost than the original walk from u to v. Hence, we have a spanning tree of at most twice the cost as F^*.

2. (7 points) Using part 1 (even if you have not proved it), deduce that MST of G_T is a 2-approx for minimum Steiner tree.

We know that $2 \times w(\text{Steiner tree}) \geq w(\text{spanning tree of part (a)}) \geq w(\text{MST})$ and we know that this MST is a feasible solution as it connects all vertices in T, so the MST is a 2-approx.

4 Randomized Algorithms (20 points)

Consider the Exact-3-SAT problem. There are n boolean variables $\{x_1, \ldots, x_n\}$. A *literal* is a variable x_i or its negation $\neg x_i$. We are given a set J of $m = |J|$ clauses; each clause is the OR of exactly three **distinct** literals on **distinct** variables, i.e., of the form $(y_1 \vee y_2 \vee y_3)$ where no variable appears twice and no clause contains both a variable and its negation (for example, clauses $(x_i \vee x_i \vee x_j)$ and $(\neg x_i \vee x_i \vee x_j)$ are not allowed). The goal is to assign truth values to maximize the number of satisfied clauses in J. A clause is satisfied if at least one of its literals is True.

We will analyze the randomized algorithm that sets each variable to be True or False independently with probability 1/2.

1. (5 points) Let p be the probability that a fixed clause $(y_1 \vee y_2 \vee y_3) \in J$ is satisfied. Compute p. (This probability does not depend on which clause we are talking about.)

   Each literal is True with probability 1/2, independently. The only way the clause is unsatisfied is if all three literals are False, which happens with probability $(1/2)^3 = 1/8$. Thus $p = 1 - 1/8 = 7/8$.

2. (5 points) What is $\mathbb{E}[S]$, the expected number of satisfied clauses? Express it using p or substitute the value computed above.

   By linearity of expectation, $\mathbb{E}[S] = \sum_{C \in J} \Pr[C \text{ is satisfied}] = pm = \frac{7}{8}m$.

3. (5 points) Let OPT be the maximum possible number of satisfied clauses. Show that the above randomized algorithm gives a p-approximation to maximizing number of satisfied clauses.

   Since OPT $\leq m$ (it can at best satisfy every clause), we have $\mathbb{E}[S] = pm \geq p\,\mathrm{OPT}$. With $p = 7/8$, this is a $(7/8)$-approximation in expectation.

4. (5 points) Using Markov's inequality, what is an upper bound on the probability that the randomized algorithm satisfies more than $11/10 \cdot \mathbb{E}[S]$ clauses?

   By Markov, this is at most $\frac{1}{\frac{11}{10}} = 10/11$.

END OF EXAM