HW 8

Due: November 6th, 2025

Problem 1: Approximation for Multicut in Trees

Consider the multicut problem in trees. We are given a tree $T = (V, E)$, a set of k terminal pairs $(s_1, t_1), \ldots, (s_k, t_k)$, and a non-negative cost $c_e \geq 0$ for each edge $e \in E$. The goal is to find a minimum-cost set of edges $F \subseteq E$ such that for all i, s_i and t_i are in different connected components of the graph $G' = (V, E - F)$.

Let P_i be the set of edges in the unique path in T between s_i and t_i. The problem can be formulated as the following integer program (IP):

$$\begin{aligned} \text{minimize} \quad & \sum_{e \in E} c_e x_e & \\ \text{subject to} \quad & \sum_{e \in P_i} x_e \geq 1, & 1 \leq i \leq k, \\ & x_e \in \{0, 1\}, & e \in E. \end{aligned}$$

We will analyze a primal-dual algorithm based on the LP relaxation of this IP.

Part A. Primal LP and its Dual

First, write the Linear Programming (LP) relaxation for the multicut IP given above. Then, write the dual of this LP. Finally, state the primal and dual complementary slackness conditions.

Part B. The Primal-Dual Algorithm

Consider the following primal-dual algorithm. We root the tree at an arbitrary vertex r. Let $depth(v)$ be the number of edges on the path from r to v. Let $lca(s_i, t_i)$ be the vertex v on the s_i-t_i path P_i whose depth is minimum.

The algorithm proceeds as follows:

- Start with a primal-infeasible solution $x_e = 0$ for all e (so $F = \emptyset$) and a dual-feasible solution $y_i = 0$ for all i.
- While the primal solution is not feasible (i.e., there exists a "violated constraint," meaning an $s_i - t_i$ path P_i that is not cut by F):
  1. Find the violated constraint j (corresponding to an un-cut path P_j) that maximizes $depth(lca(s_j, t_j))$ among all violated constraints.
  2. Increase the corresponding dual variable y_j until some dual constraint becomes tight for an edge $e \in P_j$. That is, $\sum_{i:e \in P_i} y_i = c_e$.
  3. Add every edge e that just became tight to the primal solution F (i.e., set $x_e = 1$).

Your task: Briefly explain why the final primal solution F is a feasible solution to the original multicut problem and why the final dual solution y is feasible.

Part C. The Core Lemma

The key to the approximation proof is analyzing how many edges from a single path P_i can be included in the final cut F. Let $v_i = lca(s_i, t_i)$. The path P_i is composed of two "ancestral" paths: P_i^s (from s_i to v_i) and P_i^t (from t_i to v_i).

Prove that the final cut F can contain at most one edge from P_i^s and at most one edge from P_i^t. That is, show $|P_i \cap F| \leq 2$ by proving $|P_i^s \cap F| \leq 1$ and $|P_i^t \cap F| \leq 1$.

Part D. The Approximation Ratio

Using the final primal solution F and the final dual solution y from the algorithm, you will now prove the approximation ratio. First, write the total cost of the primal solution, $Cost(F) = \sum_{e \in F} c_e$, in terms of the dual variables y_i. Next, use your result from Part (c) to bound this expression. Finally, use this bound and the relationship between the primal and dual objective values to prove that this algorithm is a 2-approximation for the multicut problem in trees.

Problem 2: Another Algorithm for Approximate Counting and Sampling of Independent Sets on Lines and Cycles

Consider the following simple Markov chain whose states are independent sets in a graph $G = (V, E)$. To compute X_{i+1} from X_i:

- choose a vertex v uniformly at random from V, and flip a fair coin
- if the flip is heads and $v \in X_i$, then $X_{i+1} = X_i \setminus \{v\}$

- if the flip is heads and $v \notin X_i$, then $X_{i+1} = X_i$
- if the flip is tails, $v \notin X_i$, and adding v to X_i still gives an independent set, then $X_{i+1} = X_i \cup \{v\}$
- if the flip is tails and $v \in X_i$ or there is an edge from some vertex in X_i to v, then $X_{i+1} = X_i$

Part A.

Show that the stationary distribution of this Markov chain is uniform over all independent sets

Part B.

We consider this Markov chain specifically on cycles and line graphs. For a line graph with n vertices, the vertices are labeled 1 to n, and there is an edge from 1 to 2, 2 to 3, $\ldots, n-1$ to n. A cycle graph on n vertices is the same with the addition of an edge from n to 1.

Propose a coupling (X_t, Y_t) for this Markov chain such that, on line graphs and cycle graphs, if $D_t = |X_t - Y_t| + |Y_t - X_t|$ is the number of vertices on which the two independent sets disagree, then at each step the coupling is at least as likely to reduce D_t as to increase D_t.

Part C.

With the coupling from Part B, argue that you can use this Markov chain to sample independent sets on a cycle graph or line graph approximately uniformly at random in polynomial time (and therefore, approximately count the number of independent sets in polynomial time).