Examen Final

1. **Tree with incentives and penalties.** AdAll is a marketing company that is planning an advertising campaign in a social network by traditional word of mouth with incentives. The social network is given by an undirected graph $G = (V, E)$, having edge cost $c_e \geq 0$, for each $e \in E$, and a penalty $p_i \geq 0$, for each $i \in V$. The values c_e, for $e = (u, v)$ represent the incentive estimated by AdAll that is required to guarantee that the ad will be passed from *u* to *v* or viceversa, assuming that it reaches *u* (or *v*). The value p_u is an estimation of the cost of excluding u from the campaign. AdAll has an initiating agent $r \in V$ and wants to estimate the suitability of the cheapest campaign. The goal is to find a tree T rooted at *r* that minimizes the cost of the edges in the tree plus the penalties of all vertices not in the tree.

   AdAll has a polynomial time algorithm (Simple) for solving a simplified version of the problem in which no vertex can be excluded from the campaign.

   In the following we use indicator variables y_i for vertex $i \in V$ and x_e for edge $e \in E$. Furthermore, we use $\delta(S)$, for $S \subseteq V$, to denote the edges crossing the cut $(S, V - S)$.

   (a). Consider that $x_e, y_i \in \{0, 1\}$ verify the following set of inequalities

   $$\sum_{e \in \delta(S)} x_e \geq y_i \qquad \forall S \subseteq V - \{r\}, S \neq \emptyset, \forall i \in S$$

   Show that in such a case the subgraph formed by the vertices with $y_i = 1$ and the edges with $x_e = 1$, minimizing the objective function $\sum_{e \in E} c_e x_e$ is a tree.

   Per a la resta del problema, assumim que el subgraf $G^* = (V^*, E^*)$ és la solució òptima al problema plantejat. Assumim també que les variables indicadores prenen els valors següents:

   $$y_i = \begin{cases} 1, & i \in V^* \\ 0, & altrament \end{cases} \qquad x_e = \begin{cases} 1, & e \in E^* \\ 0, & altrament \end{cases}$$

   L'expressió $\sum_{e \in \delta(S)} x_e \geq y_i$ ens indica que, fixades V^* i S, si hi ha algun vèrtex $i \in V^*, S$ el nombre d'arestes $e \in E^*, \delta(S)$ ha de ser major o igual a 1. Això ho podem expressar de la manera següent:

   $$S \cap V^* \neq \emptyset \Rightarrow \delta(S) \cap E^* \neq \emptyset$$

   Primer de tot, demostrarem que el subgraf G^* és connex.

Això ho demostrarem amb una reducció a l'absurd: Suposem que el graf G^* no és connex. En aquest cas, podem descompondre $V^* = V_1^* \cup V_2^* \cup \ldots \cup V_k^*$. Assumim per simplicitat que $r \in V_1^*$. Si agafem $S = V_i^*$ amb $i \neq 1$, tindrem $S \cap V^* = V_i^* \neq \emptyset$, però $\delta(S) \cap E^* = \delta(S) \cap E_i^* = \emptyset$.

Això vol dir que, si G^* no és connex, no complirà les restriccions anteriors. Com que sabem que G^* complirà les restriccions anteriors, aleshores podem concloure que G^* ha de ser connex.

A continuació, demostrarem que el subgraf G^* és acíclic.

Això ho demostrarem també amb una reducció a l'absurd: Suposem que el graf G^* té algun cicle. Si eliminem una aresta $a \in E^*$ a aquest cicle no violarem la restricció, ja que per a que això passés hauríem de tenir $S \cap V^* \neq \emptyset$, $a \in \delta(S)$ i $\delta(S) \cap E^* = \emptyset$.

Si tenim les dues primeres, però, ja no podem tenir la tercera, ja que en ser a part d'un cicle, $\delta(S)$ haurà d'incloure com a mínim una altra aresta del cicle. Així, acabarem tenint $\delta(S) \cap E^* \neq \emptyset$ igualment.

Per tant, $G^* - \{a\}$ segueix complint les restriccions anteriors. Ara assumim que tenim el cost òptim següent:

$$cost(G^*) = \sum_{e \in E} c_e x_e$$

El cost del nou graf serà el següent:

$$cost(G^* - \{a\}) = \sum_{e \in E - \{a\}} c_e x_e$$

Per tant, tenim:

$$cost(G^* - \{a\}) \leq cost(G^*)$$

Així, vol dir que per a tot graf amb un cicle podem obtenir un graf sense cicles amb un cost igual o menor. Per tant, podem concloure que G^* ha de ser acíclic.

Com que hem demostrat que G^* ha de ser un graf connex i acíclic, podem concloure que G^* és un arbre.

(b). Using the above result provide a ILP formulation of the problem. Observe that the number of equations is $n2^n$.

La formulació del problema de programació lineal entera és la següent:

$$\min_{x_e, y_i} \left\{ \sum_{e \in E} x_e c_e + \sum_{i \in V} p_i (1 - y_i) \right\}$$

Subjecte a:

$$\sum_{e \in \delta(S)} x_e \geq y_i \qquad \forall S \subseteq V - \{r\}, S \neq \emptyset, \forall i \in S$$
$$x_e, y_i \in \{0, 1\}$$

En aquesta formulació, tindrem una restricció per cada possible $S \subseteq V - \{r\}$ i per cada $i \in S$, a més d'una restricció per vèrtex i aresta. El nombre de possibles valors de S és $2^{n-1} - 1$, ja que triem d'entre només $n - 1$ vèrtex i no podem no triar-ne cap. El nombre de possibles valors de i és com a molt n per a cada tria de S. Així, en total tindrem $O(n2^n)$ restriccions.

(c). Consider the LP obtained after relaxing the ILP formulation. Let (x^*, y^*) be an optimal solution of the LP and let $\alpha \in (0, 1]$. Let $U = \{i \in V | y_i^* \geq \alpha\}$ and let T be the tree produced by SIMPLE on input $G[U]$. Show that

$$\sum_{e \in T} c_e \leq \frac{2}{\alpha} \sum_{e \in E} c_e x_e^*$$

And that

$$\sum_{i \in V - V(T)} p_i \leq \frac{1}{1 - \alpha} \sum_{i \in V} p_i (1 - y_i^*)$$

Primer, demostrarem la primera desigualtat. Per a fer-ho, demostrarem primer la proposició següent:

$$\forall e \in T \left(x_e^* \geq \frac{\alpha}{2} \right)$$

###

Aquesta demostració la farem per reducció a l'absurd. Suposem:

$$\exists e \in T \left(x_e^* < \frac{\alpha}{2} \right)$$

Si triem $S \subseteq V$ tal que $\delta(S) = e$, aleshores tenim:

$$\sum_{e \in \delta(S)} x_e^* \geq y_i$$

Com que $e \in T$, aleshores el vèrtex d'un dels extrems de l'aresta complirà $i \in V(T)$ i per tant tindrem $y_i^* \geq \alpha$. Així, tenim $x_e^* \geq \alpha$, que és una contradicció. Per tant, podem concloure:
##

$$\forall e \in T \left(x_e^* \geq \frac{\alpha}{2} \right)$$

Així, podem afirmar:

$$\sum_{e \in E} c_e x_e^* \geq \sum_{e \in T} c_e x_e^* \geq \sum_{e \in T} c_e \frac{\alpha}{2} = \frac{\alpha}{2} \sum_{e \in T} c_e$$

D'aquí, podem aïllar:

$$\sum_{e \in T} c_e \leq \frac{2}{\alpha} \sum_{e \in E} c_e x_e^*$$

Ara, demostrarem la segona desigualtat. Ara, cal tenir en compte que $y_i^* \geq \alpha \Leftrightarrow i \in V(T)$. Així, podem afirmar que $i \notin V(T) \Rightarrow y_i^* \leq \alpha$.

Així, podem afirmar:

$$\begin{aligned} \sum_{i \in V} p_i(1 - y_i^*) &= \sum_{i \in V(T)} p_i(1 - y_i^*) + \sum_{i \in V - V(T)} p_i(1 - y_i^*) \\ &\geq \sum_{i \in V - V(T)} p_i(1 - y_i^*) \geq \sum_{i \in V - V(T)} p_i(1 - \alpha) \end{aligned}$$

D'aquí, podem aïllar:

$$\sum_{i \in V - V(T)} p_i \leq \frac{1}{1 - \alpha} \sum_{i \in V} p_i(1 - y_i^*)$$

(d). Assuming that a solution to the LP can be computed in polynomial time, can you design a 3-approximation for AdAll's problem?

Si realitzem tots els passos indicats als apartats anteriors, tenim que aquest algorisme computa una aproximació de la forma:

$$A(G) = \sum_{e \in T} c_e + \sum_{i \in V - V(T)} p_i$$

D'altra banda, tenim:

$$opt(G) = \sum_{e \in E} c_e x_e^* + \sum_{i \in V} p_i (1 - y_i^*)$$

Aplicant les desigualtats de l'apartat (c), tenim:

$$A(G) \leq \frac{2}{\alpha} \sum_{e \in E} c_e x_e^* + \frac{1}{1 - \alpha} \sum_{i \in V} p_i (1 - y_i^*)$$

Si fixem $\alpha = \frac{2}{3}$ tenim:

$$A(G) \leq 3 \sum_{e \in E} c_e x_e^* + 3 \sum_{i \in V} p_i (1 - y_i^*) = 3 \left(\sum_{e \in E} c_e x_e^* + \sum_{i \in V} p_i (1 - y_i^*) \right)$$

Per tant, podem concloure:

$$A(G) \leq 3 \cdot opt(G)$$

Per tant, l'algoritme proposat és una 3-aproximació del problema.