Assignment 2

Due date: January 11th, beginning of the lecture.

Abstract

There are three parts that are (almost) completely independent. Harder questions are denoted by (*). Questions are in a "logical" order which means that the complexity is not necessarily increasing ! You usually does not need to solve done question i to solve question $i+1$. So if you don't solve a question, do not hesitate to look at the next ones !

1 Identification on interval graphs

Let $G = (V, E)$ be a graph. Let $v \in V$, we denote by $N(v)$ the set of neighbors of v and $N[v]$ the set of neighbors of v plus v. A set $S \subseteq V$ is *identifying* the graph if:

- S is a *dominating set*, i.e., for every $v \in V$, $v \in S$ or there exists $s \in S$ such that (s, v) is an edge.
- For every pair of vertices $u, v \in V$, $N[u] \cap S \neq N[v] \cap S$.

The goal of this part consists in proving that the identification problem (which consists in finding an identifying set of minimal size) has an approximation algorithm for interval graphs. Given two sets X and Y, we denote by $X \Delta Y$ the symmetric difference of X and Y. In other words, $X \Delta Y = X \setminus Y \cup Y \setminus X$.

Question 1. Formulate the identitication problem as an ILP (denoted by P in what follows).

Question 2.
We now consider the two following ILP:

$$P_{inter}$$
$$\min \sum_{i \in V} x_i$$
$$\forall \boldsymbol{jk} \in \boldsymbol{E} \quad \sum_{i \in N[j] \triangledown N[k]} x_i \geq 1$$
$$x_i \in \{0,1\} \quad \forall i \in V$$

$$P_{disj}$$
$$\min \sum_{i \in V} x_i$$
$$\forall \boldsymbol{jk} \notin \boldsymbol{E} \quad \sum_{i \in N[j] \triangledown N[k]} x_i \geq 1$$
$$\sum_{i \in N[j]} x_i \geq 1 \quad \forall j \in V \; x_i \in \{0,1\} \quad \forall i \in V$$

Question 2.a. Prove that $OPT(P_{inter}) \leq OPT(P)$ and $OPT(P_{disj}) \leq OPT(P)$.

Question 2.b. Prove the union of the solutions of the ILPs P_{inter} and P_{disj} is identifying the graph.

Question 3.
An *interval graph* is an intersection of (closed) segments in the plane. In other words,

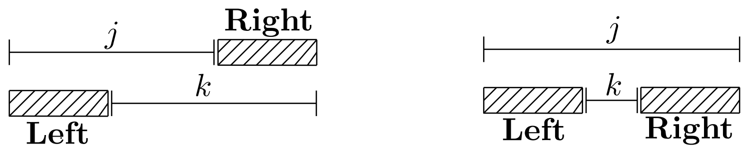

Figure 1: Given two intersecting intervals j and k, one can construct two areas **Left** and **Right** partitioning $N[j] \bigtriangledown N[k]$ between L_{jk} the set of intervals that end in **Left**, and R_{jk} the set of intervals that begin in **Right**. This figure shows how to find **Left** and **Right** depending on the configuration of j and k: either one is included in the other, or not.

every vertex is associated to an interval of the real line. (We assume that every vertex is given with its corresponding interval). And there is an edge between two vertices if and only if their corresponding intervals intersect. We put an arbitrary order on the real line. The *begin date (resp. stopping date) of an interval* x is the first (resp. last) point p of the real line (in the order) such that $p \in x$.

Our goal is to prove that we can approximate the solutions of P_{inter} and P_{disj} in interval graphs. Let us first start by considering the ILP P_{disj}. Let $x_1, \ldots, x_n$ be the vertices of G sorted by stopping date. Consider the following algorithm.

$X = \emptyset$
while $G \neq \emptyset$ **do**
  Let x be the vertices of G with the smallest stopping time.
  $x \leftarrow X \cup \{x\}$.
  Delete all the vertices of $N[X]$ from G.
end while

Question 3.a. Prove that X is a maximum independent set of G.

Question 3.b. Prove that X is a solution of P_{disj}.

Question 3.c. Prove that $|X| \leq 2 \cdot OPT(P_{disj})$.

Question 4.
Le $x^* = (x_1^*, \ldots, x_n^*)$ be an optimal solution of P_{inter}^*. For every $jk \in E$, Figure 1 shows how to partition $N[j] \bigtriangledown N[k]$ into two parts L_{jk} (stands for Left) and R_{jk} (stands for Right). The set L_{jk} is composed of the intervals that end between the begin dates of j and k, and R_{jk} is composed of the intervals that begin between the end dates of j and k. L_{jk} and R_{jk} are obviously disjoint subsets.

Let us now define two subsets of vertices L and R as follows:

$$L = \left\{ jk \in E \;\middle|\; \sum_{i \in L_{jk}} x_i^* \geq \frac{1}{2} \right\} \qquad \text{and} \qquad R = \left\{ jk \in E \;\middle|\; \sum_{i \in R_{jk}} x_i^* \geq \frac{1}{2} \right\}$$

We (again !) define the following two integer linear programs:

$$P_L$$
$$\min \sum_{i \in V} x_i$$
$$\forall \boldsymbol{jk} \in \boldsymbol{L} \sum_{i \in L_{jk}} x_i \geq 1$$
$$x_i \in \{0,1\} \quad \forall i \in V$$

$$P_R$$
$$\min \sum_{i \in V} x_i$$
$$\forall \boldsymbol{jk} \in \boldsymbol{R} \sum_{i \in R_{jk}} x_i \geq 1$$
$$x_i \in \{0,1\} \quad \forall i \in V$$

Question 4.a. Prove that the union of a solution of P_L and of a solution of P_R is a solution of P_{inter}.

Question 4.b. An *interval matrix* is a $\{0,1\}$-matrix where, on each row, all the 1 coefficients are consecutive. In other words, if for all i and $j \leq k \leq \ell$ such that $a_{i,j} = a_{i,\ell} = 1$ implies $a_{i,k} = 1$.
Prove that the constraint matrices of both P_L and R_R are interval matrices.

Question 4.c. Using the fact that interval matrices are TU, what can you deduce?

Question 4.d. Deduce that you can find in polynomial time a solution of P_{inter} of size at most $4 \cdot OPT(P_{inter})$.

Question 5. Deduce a 6-approximation algorithm running in polynomial time for identification on interval graphs.

Collaboration Policy

You may discuss anything related to this project with your classmates. However, you must write your own solution; you are not allowed to copy-paste **anything** from your classmates or from the internet.