Problem 23. Consider a set $X = \{x_1, \ldots, x_n\}$ and a collection $Y_1, Y_2, \ldots, Y_m$ of subsets of X (i.e. $Y_i \subseteq X$ for each i). We say that a set $S \subseteq X$ is a *Special Set* for the collection $Y_1, Y_2, \ldots, Y_m$ if S contains at least one element from each Y_i (in other words, $S \cap Y_i \neq \emptyset$ for each i).

We now define the *Special Set Problem* as follows. We are given a set $X = \{x_1, \ldots, x_n\}$, a collection $Y_1, Y_2, \ldots, Y_m$ of subsets of X, and a positive integer k. We are asked: Is there a *special set* $S \subseteq X$ for $Y_1, Y_2, \ldots, Y_m$ so that the size of S is at most k?

We want to show that *Special Set Problem* is NP-complete.

(a) (3 pts) Prove that *Special Set Problem* is NP.

(b) (12 pts) Prove that *Special Set Problem* is NP-hard. (Hint: You may use *Vertex Cover* for reduction).

sol. This problem is originally known as *Hitting Set Problem.* We slightly change the name of problem to prevent the searching on online.

(a) (No partial points) For a given certificate S, we can verify in polynomial time whether the size of S is at most k and S intersects each of the sets $Y_1, Y_2, \cdots, Y_m$, since checking the intersection of two sets takes at most $O(n^2)$.

(b) (***Solution 1***) Given an arbitrary graph $G = (V, E)$, we construct an instance of *Special Set* as follows: First, let X be a vertex set of G, i.e. $X = V(G)$. And for each edge $e = (u, v) \in G$, we construct a set $Y_e = \{u, v\}$, then we have $|E(G)| = m$ sets. We now claim that G has a vertex cover of size k if and only if $\{Y_e\}_{e \in E(G)}$ has a *Special Set* $S \subseteq X$ of size at most k. (9 points)

($\Rightarrow$) Suppose G has a vertex cover VC of size at most k. Then, for each edge $e = (u, v)$, either $u \in VC$ or $v \in VC$, in other words, $Y_e \cap VC \neq \emptyset$. Therefore, $VC = S$ is a *Special Set* of size at most k.

($\Leftarrow$) Conversely, suppose S is a *Special Set* of size at most k for the collection of sets $\{Y_e\}_{e \in E(G)}$. Then, for each edge $e \in E(G)$, at least one endpoint will be covered by S which means S is a vertex cover of G with size at most k. (3 points)

(***Solution 2***) Given an arbitrary graph $G = (V, E)$, we construct an instance of *Special Set* as follows: First of all, let X be a vertex set of G, i.e. $X = V(G)$. And, for each vertex $v \in V(G)$, we construct a set Y_v which contains v and all neighbors of v in G. We now claim that G has a vertex cover of size k if and only if $\{Y_v\}_{v \in V(G)}$ has a *Special Set* $S \subseteq X$ of size at most k. (9 points)

($\Rightarrow$) Suppose G has a vertex cover VC of size at most k. Then, for each Y_v, at least one element in VC must be contained in Y_v by the definition of vertex cover. In other words, $Y_v \cap VC \neq \emptyset$. Therefore, $VC = S$ is a *Special Set* of size at most k.

($\Leftarrow$) Conversely, suppose S is a *Special Set* of size at most k for the collection of sets $\{Y_v\}_{v \in V(G)}$. Then, for each vertex $v \in V(G)$, we can see either $v \in S$ or one of neighbors of v is in S. Therefore, S is also a vertex cover of G with size at most k. (3 points)

(***Wrong Solution 1***) (0 point) Many students tried to construct the graph based on the sets X and $Y_1, \ldots, Y_n$. However, to show the reduction from vertex cover, we need to show the opposite way, which means given an arbitrary graph, we need to construct the sets X and $Y_1, \ldots, Y_n$.

(***Wrong Solution 2***) (0 point) Some students tried to construct the sets X and $Y_1, \ldots, Y_n$ based on the result of vertex cover. However, to show the reduction from vertex cover, we need to construct the sets first.