Vivek Pandey

vpandey35@gatech.edu

Problem 2 (Hitting Set)

References:

Joves Notes

Approach:

Given a family of sets {S_1, S_2, ..., S_n} and a budget b > 0, the Hitting Set problem returns a set H of size less than or equal to **b** which intersects every S_i, if such an H exists.

In order to prove that this Hitting Set problem is NP-complete, I will show that

- A solution to Hitting Set problem can be verified in polynomial time such that $Hitting\ Set\ \in NP$
- A known NP-complete problem, the Vertex Cover (VC) problem, can be reduced to a Hitting Set problem i.e., $Vertex\ Cover\ \rightarrow\ Hitting\ Set$

NP Proof:

For given family of input sets {S_1, S_2, ..., S_n} and budget b to the Hitting Set problem, we look for a solution set H.

- If a solution doesn't exist, we return NO.
- If a solution set H does exist, we need to verify H is valid solution by checking the following two things:
  - |H| <= b. O(n) time complexity

- H contains at least one element from each input sets {S_1, S_2, ..., S_n} which can be done in polynomial time even if we run an optimal nested loop to check this.

Since we can verify the solution of Hitting Set problem in polynomial time, we have proven that $Hitting\ Set\ \in NP$

Input: $(Vertex\ Cover\ \rightarrow\ Hitting\ Set)$

We take the input to the Vertex Cover problem - an undirected graph G = (V, E) and integer b dictating the maximum size of vertex cover. Now we take all the edges in G such that each edge $e_i = (u, v) \in E$ is turned into an input set $S_i = \{\ u, v\}$ to the Hitting Set Problem. This conversion of edges to sets can be done in polynomial time of O(|E|) time. The maximum vertex cover size **b** can be passed to the Hitting Set Problem as the budget **b.** Hence, the overall runtime remains polynomial in original input size.

Output: $(Vertex\ Cover\ \rightarrow\ Hitting\ Set)$

If Hitting Set returns, NO, return NO as the solution for Vertex Cover problem. Otherwise, Hitting Set returns a set H of size less than or equal to **b** which intersects every S_i i.e., H contains at least one of the endpoints of every edge $e_i = (u, v) \in E$. Therefore, the solution of Hitting Set problem can be returned as-is as the solution of Vertex cover. Therefore, this output conversion has an O(1) time complexity.

Correctness:

We can show the correctness of this reduction by showing that:

$$Solution\ to\ Vertex\ Cover\ exists\ \ \Leftrightarrow Solution\ to\ Hitting\ Set\ exists$$

First, we show the forward implication i.e., if solution to vertex cover exists, then solution to hitting set exists. For a vertex cover solution to exist, at least one of the vertices in each edge need to exist in the solution set of vertices. And if this condition is fulfilled, upon drawing the analogy that each edge is just a set of its endpoint vertices, it is guaranteed at least one member of a set/edge is guaranteed to be in the solution of hitting set.

Next, we show the backward implication i.e., if solution to hitting set exists, then solution to vertex cover exists. For a hitting set solution to exist, at least one item from each input set needs to appear in the final solution. And if this condition is met, then upon drawing the analogy that each edge is just a set of its endpoint vertices, it is guaranteed that for every edge $e_i = (u, v) \in E$, either u or v will belong to the final solution of vertex cover.

Therefore, finding a hitting set of size b in the family of sets is the same as finding the vertex cover of size b in the family of edges (treated as edge-vertices sets) of a graph. So if we can find solution to one, we are guaranteed to find solution for the other.