CS5820 HW-8

Renhao Lu, NetID: rl839

2. Monotone SAT is a NP-complete

1. Monotone SAT is in NP

Certificate: the value of $x_1, x_2, x_3, \dots, x_n$

Check: if $\Phi = True$ and the the false number in $\{x_1, x_2 \dots x_n\}$ is no less than k, the answer is yes. Otherwise, answer is no.

2. Reduction of independent set problem to Monotone SAT problem:

(1) Convert input

Assume the input of independent set problem is integer k and a graph $G = (V, E)$, where $V = \{1,2,..,n\}$ and $E = \{(u,v)|u,v \in V\}$, $|E| = m$. The independent set problem would be whether we can find an independent set $E' \in E$ and $|E'| = k$.

Then we create n variables $x_1, x_2, \dots, x_n$, each x_i is corresponding to node i. Then $\Phi = (x_{u1} \vee x_{v1}) \wedge (x_{u2} \vee x_{v2}) \vee \dots \vee (x_{um} \vee x_{vm})$, where each clause $x_{uj} \vee x_{vj}$is corresponding to an edge $(u_j, v_j) \in E$. Then $x_1, x_2, \dots, x_n$, Φ, and integer $k' = k$ is the input of Monotone SAT problem. The monotone SAT problem would be whether we can find a solution of $x_1, x_2, \dots, x_n$, so that Φ is true and the number of false in $\{x_i\}$ is not less than k.

(2) Convert output

The yes/no answer of monotone SAT would be the yes/no answer of independent set problem

(3) Running time

Running time for converting: Creating n variables takes $O(n)$, and creating Φ takes $O(m)$. So total converting time would be $O(n + m)$. Because $m \le \frac{n(n-1)}{2}$, this is in polynomial time.

So if monotone SAT can be solved in polynomial time, then independent set can also be solved in polynomial time.

(4) Proof of correctness

We claim that the result of monotone SAT problem is equivalent to equivalent to independent set problem.

If monotone SAT has a solution, then we can create the independent set $V' = \{i | \forall x_i \; is \; false\}$, and $|V'| = k' = k$. We can claim that two node in V' are connect in G because otherwise one clause $(x_i, x_i{}')$ would be false. Hence the V' is a solution of independent set problem.

If independent set problem has a solution, we assume the solution is V', $|V'| = k$. Then for any x_i that $i \in V'$, we set x_i is false; for other x_i, we set them as true. Then the false number in $\{x\}$ is k, and Φ must be true, because no two false x_i should be in one clause. If two false x_i are in one clause, it means two nodes in V' are connected. Hence $\{x_i\}$ is a solution of monotone SAT.

3. Sparse Maximum Flow

1. Sparse maximum flow is in NP

Certificate: The flow value of each edge in $G = (V, E)$

Check: if there are at most k non-zero edge in E, then the answer is yes, otherwise is no.

2. Reduction of vertex cover problem to sparse maximum flow problem

(1) Convert input

Given the input of vertex cover problem: undirected graph $G = (V, E)$ and integer k (assume $|V| = n, |E| = m$). Create a directed graph $G' = (V', E')$ and integer k' for input of sparse maximum flow problem. Here in V', we split every node i in input graph V into two nodes i_s and i_t in V', where i_s is connect to the source s with capacity of ∞ and i_t is connect to the sink t with capacity 1. For each edge (u, v) from the vertex cover input E, we create two edges (u_s, v_t) and (v_s, u_t) in E' with capacity 1. We also need to add edges (i_s, i_t) for each i with capacity 1.

In other words,

$$V' = \{s, t\} \cup \{i_s, i_t | \forall i \in V\}$$

$$E' = \{(s, i_s)\} \cup \{(\quad)\} \cup \{(i_s, i_t)\} \cup \{(u_s, v_t), (v_s, u_t) | \forall (u, v) \in E\}$$

$$c_e = \begin{cases} +\infty, & e = (s, i_s) \\ 1, & e = (i_t, t) \\ 1, & e = (u_s, v_t) \ or \ (v_s, u_t) \\ 1, & e = (i_s, i_t) \end{cases}$$

$$k' = 2n + k$$

(2) Convert output

If each edge (s, i_s) that has a non-zero flow, we pick the node i in the vertex set for the solution of the vertex cover problem. If the sparse maximum flow problem has a solution, the vertex cover problem has a solution. Otherwise, the vertex cover problem doesn't have a solution.

(3) Running time

Converting input: in new graph G', creating V' takes $O(2 + 2n) = O(n)$, creating E' takes $O(n + n + 2m) = O(n + m)$, creating c_e takes $O(n + m)$. So totally $O(n + m)$. Because $m \leq \frac{n(n-1)}{2}$, $O(n + m)$ is polynomial.

(4) Proof of correctness

We claim that the result of sparse maximum flow problem is equivalent to equivalent to vertex cover problem.

Claim 1: the maximum flow in G' is n

Proof: Consider one $s - t$ cut (A_0, B_0), where $B_0 = \{t\}$, $A_0 = V'/B_0$, then $c(A_0, B_0) = n$ because there totally n (i_t, t) edges with capacity 1. Then we know that the max flow $f_{max} \leq c(A_0, B_0) = n$.

Next, we consider a flow:

$$s - 1_s - 1_t - t, flow = 1$$

$$s - 2_s - 2_t - t, flow = 1$$

$$s - 2_s - 2_t - t, flow = 1$$

$$s - 3_s - 3_t - t, flow = 1$$

...

$$s - n_s - n_t - t, flow = 1$$

This is a valid flow and obviously the total flow value is n. Hence the $f_{max} \geq n$. As a summary, $f_{max} = n$. ■

Claim 2: The number of edge (s, i_s) with non-zero flow value is at most $k' - 2n = k$

Proof: In the solution of sparse maximum flow problem, the total flow value should be the maximum flow value, which is n according to claim 1. Hence, each (i_t, t) has a flow of 1. Meanwhile, for each i_t, at least one edge in $\{(i_s, i_t)\} \cup \{(u_s, i_t)\}$ should have a nonzero flow according to the flow conservation constraint. So far, the total nonzero edge $\geq 2n$. Hence the number of non-zero (s, i_s) is at most $k' - 2n = k$.

If sparse maximum flow problem has a solution, we pick each node i into the vertex set V_0 if edge (s, i_s) has a non-zero flow value. According to claim 2, the size of $|V_0| \leq k$. Because the maximum flow is n, for each i_t, there must be at least one non-zero flow from $i_s \in V_0$. Hence the nodes in V_0 can cover all vertexes. Hence V_0 is a solution for vertex cover problem.

If vertex cover problem has a solution, we note this solution is a node set V_0. Then we create a flow in G' according to V_0. For all the node $i \notin V_0$, we set the flow of (s, i_s) as 0. Then we set the flow of all (i_t, t) as 1. Because V_0 is a solution of vertex cover problem, for each edge $(u, v) \in E$, at least one node u or v should be in V_0. If $u \in V_0, v \in V_0$, we set (u_s, u_t) and (v_s, v_t) flow as 1. If $u \in V_0, v \notin V_0$, we set (u_s, v_t) and (u_s, u_t) flow as 1. If $u \notin V_0, v \in V_0$, we set (v_s, u_t) and (v_s, v_t) flow as 1. Then we set all (s, i_s) according to flow conservation constrain and all other edges in G' with zero flow. The total flow is equal to the sum of all (i_t, t), which is n, hence it is a maximum flow. All i_t has exactly 1 inflow and 1 outflow, and we set flow of i_s according to flow conservation constrain, so the flow is also a valid flow. Meanwhile, the nonzero flow edges including all (i_t, t), (u_s, v_t), (s, i_s) that are described above. $|\{(i_t, t)\}| = \left|\{(u_s, v_t)_{f \neq 0}\}\right| = |V| = n$, and $\left|\{(s, i_s)_{f \neq 0}\}\right| = |V_0| = k$. So the total nonzero flow edge is $2n + k = k'$. Hence the flow is a solution of sparse maximum flow problem. ■