

**Theorem 1.** *Independent set is  $\mathcal{NP}$ -complete.*

**Proof.** It is easy to see that independent set is in  $\mathcal{NP}$  (convince yourself of that by constructing an appropriate verifier!), and we focus on showing  $\mathcal{NP}$ -hardness.

Specifically, we show a Karp reduction of 3SAT to IND. That is, given as input a 3CNF formula  $\Phi$ , we construct in polynomial time  $(G, k)$  such that  $\Phi$  is satisfiable if and only if  $G$  has an independent set of size  $k$ .

Let  $C_1, \dots, C_m$  be the clauses of  $\Phi$ . The graph  $G$  has  $3m$  nodes, where each triplet  $i \in [m]$  of nodes corresponds to a clause  $C_i$ . Each node in a triplet  $C_i$  is labeled by the corresponding literal  $x_j$  or  $\neg x_j$ . We add two types of edges:

- **Triplet edges:** The three nodes in each triplet are connected by three edges.
- **Consistency edges:** For each pair of nodes labeled by  $x_j$  and  $\neg x_j$ , we add an edge between the them.

We output this graph and  $k = m$ . Convince yourself that the reduction can be implemented in polynomial time. Turning to correctness,

- If  $\Phi$  is satisfiable, there is  $x$  such that  $\Phi(x) = 1$ . Since  $\Phi$  is an AND of ORs, for each  $i \in [m]$  there is at least one literal in  $C_i$  that evaluates to 1. We build an independent set of size  $m$  by adding to  $S$ , for each  $C_i$ , exactly one literal in  $C_i$  that evaluates to 1. There are  $m$  clauses, hence  $|S| = m$ .

To see that  $S$  is independent, let  $G_S$  be the induced subgraph of  $S$ , and note that: (a) We only added one node from each triplet, so no triplet-edge is included in  $G_S$ ; (b) There cannot be two nodes in  $S$  labeled by  $x_j$  and  $\neg x_j$ , because we only add a node to  $S$  if the literal evaluates to 1 in  $x$  (and for each  $j$  either  $x_j$  or  $\neg x_j$  evaluate to 1), hence there are no consistency edges in  $G_S$ .

- In the other direction, we show that if there's an independent set  $S$  of size  $m$  in  $G$  then  $\Phi$  is satisfiable.

Let  $S$  be an independent set of size  $m$ , and note that  $S$  cannot include two nodes from the same triplet (because of the triplet-edges). Since  $|S| = m$ , there is exactly one node in  $S$  from each triplet. Also, there can't be two nodes labeled by  $x_j$  and  $\neg x_j$  in  $S$  (because of consistency-edges), hence for each variable  $x_j$ , either a node  $x_j$  is in  $S$  or a node  $\neg x_j$  is in  $S$  (or neither is in  $S$ ).

We build a satisfying assignment for  $\Phi$  by setting, for each  $j$ , either  $x_j = 1$  (if a node labeled by  $x_j$  is in  $S$ ) or  $x_j = 0$  (otherwise). This is a well-defined setting of the  $x_j$ 's, by the above.<sup>1</sup> For each clause  $C_i$ , there is some node in the corresponding triplet that is included in  $S$ , and we set  $x_j$  so that the literal in  $C_i$  will evaluate to 1. Hence, each  $C_i$  is satisfied by  $x$ , meaning that  $\Phi(x) = 1$ .

Thus,  $\Phi$  is satisfiable if and only if there's an independent set of size  $m$  in  $G$ . ■

<sup>1</sup>That is, since  $S$  can't contain a node labeled by  $x_j$  and another node labeled by  $\neg x_j$ .