

## Problem Set: NP-Complete Reductions

1. [easy] Consider the two problems:

DOUBLE-SAT: Given as input a boolean formula  $\psi$ , does  $\psi$  have at least two satisfying assignments?

SAT: Given as input a boolean formula  $\psi$ , does  $\psi$  have a satisfying assignment?

Show that DOUBLE-SAT is NP-complete using a reduction from SAT.

Solution:

- a. DOUBLE-SAT  $\in$  NP.

We guess a certificate as a truth assignment to the variables  $\mathbf{x}=\{x_1, \dots, x_n\}$ . We can verify if  $\psi$  is true in poly-time, since  $\psi$  has finite number of clauses.

We define a reduction from SAT to DOUBLE-SAT as follows. For any SAT formula  $\psi$  we create a DOUBLE-SAT formula  $\psi \wedge (y \vee \neg y)$ . This trivially is done in poly-time.

We now prove that we have a *yes* SAT instance iff we have a *yes* DOUBLE-SAT instance.

$\Rightarrow$  If  $\psi$  is true (we have a *yes* SAT instance) then we must have 2 *yes* DOUBLE-SAT instances, one with  $y$  being true and one with  $y$  being false.

$\Leftarrow$  If we have 2 *yes* DOUBLE-SAT instances, one with  $y$  being true and one with  $y$  being false, this is only possible if  $\psi$  is true (we have a *yes* SAT instance).

2. [easy] Consider the two problems:

Weighted vertex cover (WVC): given a Graph  $G=(V,E)$ , a positive integer weight  $w(v)$  for each vertex  $v \in V$ , and an integer  $k$ , does  $G$  have a vertex cover of total weight at most  $k$ ?

Vertex cover (VC): given a Graph  $G=(V,E)$ , and an integer  $k$ , does  $G$  have a vertex cover of size at most  $k$ ?

Show that WVC is NP-complete by reduction from VC.

We restrict instances of WVC to have all weights as 1. This means that we have exactly the VC problem. Hence WVC is NP-complete, since WVC is at least as hard as VC, since WVC can have arbitrary positive weights.

3. [easy] Consider the following two problems:

CLIQUE: given an undirected graph  $G(V,E)$ , and an integer  $k>0$ , does  $G$  contain a clique of size  $k$ ?

HALF-CLIQUE: given an undirected graph  $G(V,E)$ , and an integer  $k>0$ , does  $G$  contain a clique of size  $k=|V|/2$ ?

Show that CLIQUE is NP-complete using a reduction from HALF-CLIQUE.

**Solution:** Clearly Half-CLIQUE  $\in$  NP, since an oracle can guess a certificate consisting of a set of  $k=|V|/2$  vertices, and we can test if these vertices are fully-connected in time  $O(V^2)$ .

To show that CLIQUE is NP-Complete, we can restrict the instances of our target problem, CLIQUE, to cases where  $k=|V|/2$  vertices. In this case CLIQUE is identical to HALF-

CLIQUE. Hence it is now trivial to show that yes instances of CLIQUE occur iff we have yes instances of HALF-CLIQUE.

#### OTHER DIRECTION OF REDUCTION

**Show that HALF-CLIQUE is NP-complete using a reduction from CLIQUE.**

We do this by showing it is in NP, and then showing the necessary reduction.

**HALF-CLIQUE is in NP**, since given a collection of  $\frac{1}{2}|V|$  vertices we can check if they are fully connected in time  $O(|V|^2)$ .

**Reduction:** We use CLIQUE for our reduction: We assume an instance  $(G,k)$  of CLIQUE and an instance  $H$  of HALF-CLIQUE.

We now define a map from pairs  $(G,k)$  to graphs  $H$  such that  $G$  has a clique of size  $k$  iff  $H$  has a clique of size  $\frac{1}{2}n$  (where  $n$ , an even number, is the number of nodes in  $H$ ). We don't want to delete anything from  $G$  as we make  $H$ , because any part of  $G$  might be important in forming a clique of size  $k$ . So we will make  $H$  by adding nodes to  $G$ . Let  $m$  be the number of nodes in  $G$ .

If  $k \geq \frac{1}{2}m$ , we can make  $H$  equal to  $G$  plus some number  $t$  of new nodes of degree 0. Then  $G$  has a clique of size  $k$  iff  $H$  has a clique of size  $k$ . We get the desired reduction by making the size of  $H$ ,  $n = m+t$ , exactly  $2k$ , so that  $k = \frac{1}{2}n$ . This can be done by making  $t = 2k - m$ , which is possible whenever  $k \geq \frac{1}{2}m$ .

If  $k < \frac{1}{2}m$  we need a different construction. Again we add  $t$  new nodes, but now we also add edges from each new node to each other node in the graph, both new and node. Thus any  $k$ -clique in  $G$  combines with the  $t$  new nodes to make a  $(k+t)$ -clique in  $H$ . (And a  $k+t$  size clique in  $H$  must include at least  $k$  old nodes, which form a clique in  $G$ .) So we need only pick  $t$  such that  $k+t = (m+t)/2$ , or  $t = m-2k$ , so that the cliques we seek in  $H$  are exactly size  $n/2$ .

4. [medium] Consider the following two problems:

CLIQUE: given an undirected graph  $G(V,E)$ , and an integer  $k > 0$ , does  $G$  contain a clique of size  $k$ ?

HALF-CLIQUE: given an undirected graph  $G(V,E)$ , and an integer  $k > 0$ , does  $G$  contain a clique of size  $k = |V|/2$ ?

Show that HALF-CLIQUE is NP-complete using a reduction from CLIQUE.

#### **Solution:**

Clearly HALF-CLIQUE  $\in$  NP, since given a certificate of a subset of vertices we can verify whether they form a clique in time polynomial in  $V$  and  $E$ .

To show that HALF-CLIQUE is NP-Complete, we give a reduction from CLIQUE to MIN-CLIQUE, as follows:

On input  $[G, k]$ , where  $G$  is a graph on  $n$  vertices and  $k$  is an integer:

1. If  $k = \lceil n \rceil / 2$ , then output  $G$ .

2. If  $k < \lceil n \rceil / 2$ , then construct a new graph  $G'$  by adding a complete graph with  $n - 2k$  vertices and connecting them to all vertices in  $G$ , and output  $G'$ .
3. If  $k > n/2$ , then construct a new graph  $G^*$  by adding  $2k - n$  isolated vertices to  $G$ , and output  $G^*$ .

We now show that our reduction is correct for each of these 3 cases.

1. When  $k = \lceil n \rceil / 2$ , it is clear that  $[G, n/2] \in \text{CLIQUE}$  if and only if  $[G] \in \text{MIN - CLIQUE}$ .
  2. When  $k < \lceil n \rceil / 2$ , if  $G$  has a  $k$ -clique, then  $G'$  has a clique of size  $k + (n - 2k) = (2n - 2k)/2$ , and therefore  $[G'] \in \text{HALF-CLIQUE}$  as  $G'$  is a graph with  $2n - 2k$  vertices. Conversely, if  $[G'] \in \text{MIN - CLIQUE}$ , that is, if  $G'$  has a clique of size  $n - k = k + (n - 2k)$ , then at most  $n - 2k$  of the clique come from the  $n - 2k$  new vertices. Therefore the remaining at least  $k$  vertices form a clique in  $G$ , and hence  $[G, k] \in \text{CLIQUE}$ .
  3. When  $k > n/2$ , if  $G$  has a  $k$ -clique, then  $G^*$  has a clique size  $k = 2k/2$ , and therefore  $[G^*] \in \text{HALF-CLIQUE}$  as  $G^*$  is a graph with  $n + 2k - n = 2k$  vertices. Conversely, if  $[G^*] \in \text{MIN - CLIQUE}$ , that is, if  $G^*$  has a clique of size  $k$ , then the clique does not contain any of the new vertices as they are isolated. Thus the clique is a  $k$ -clique of  $G$ , and hence  $[G, k] \in \text{CLIQUE}$ .
5. [easy] Consider the following two problems:

KNAPSACK: Given a set  $T$  of items, each with integer value  $v$  and integer weight  $w$ , and integers  $W$  and  $V$ , is there a subset of  $T$  of weight no more than  $W$ , with total value at least  $V$ ?

SUBSET-SUM: given a collection  $S = \{s_1, \dots, s_n\}$  of integers and a target integer  $t$ , does a sub-collection of  $S$  sum to  $t$ ?

Show that KNAPSACK is NP-complete using a reduction from SUBSET-SUM.

We restrict instances of KNAPSACK to have  $W=V$ , i.e., all values equal weights. In this case KNAPSACK is exactly the SUBSET-SUM problem, and hence is NP-complete.

6. [medium] Consider the following problems:

k-colouring: given a graph  $G(V, E)$  and an integer  $k < |V|$ , can you colour the vertices  $V$  with  $k$  colours such that no two adjacent vertices have the same colour?

Class scheduling: Given a set  $n$  of classes, each with a time slot  $S$ , and an integer  $k < n$ , can you assign the classes to  $k$  rooms such that no class has an overlapping time slot?

Show that Class scheduling is NP-complete using a reduction from  $k$ -colouring.

**Solution**: Clearly Class scheduling  $\in$  NP, since an oracle can guess a certificate consisting of a set of classes, time slots and rooms, and we can test if these classes can be scheduled in time polynomial in  $n$ .

We now show a reduction from  $k$ -colouring,  $[G(V, E), k]$  to scheduling, creating a scheduling graph  $[G'(V', E'), k]$ . For each vertex we create a class. For each overlapping time slot between vertices  $u$  and  $v$  we create an edge  $(u, v)$ , which denotes that  $u$  and  $v$  cannot be scheduled without an overlap. For each colour in  $k$ -colouring we create a classroom.

$\Rightarrow$  Suppose  $G$  has a  $k$ -colouring. By our construction this means that  $G'$  has a schedule of size  $k$ , since no classes with time-conflicts are scheduled in the same room.

$\Leftarrow$  Suppose that  $G'$  has a schedule of size  $k$ . By our construction, this can only occur if  $G$  has a  $k$ -colouring.

7. [medium] Consider the following two problems:

Multi-Path: given a weighted undirected graph  $G(V,E)$ , where loops and multiple edges are allowed and the edge weights are positive integers, a specified node  $v$ , and an integer  $k > 0$ , does there exist a path in  $G$  that starts at vertex  $v$ , never repeats an edge, and returns to  $v$  after travelling a distance of exactly  $k$ ?

SUBSET-SUM: given a collection  $S = \{s_1, \dots, s_n\}$  of integers and a target integer  $t$ , does a sub-collection of  $S$  sum to  $t$ ?

Show that Multi-Path is NP-complete using a reduction from SUBSET-SUM.

We first show that MP is in NP. A certificate for this problem is the sequence of edges in the path, which has size  $O(|E|)$ . It is straightforward to check in polynomial time that no edge is repeated, that the start and end of the path is the same, and that the total weight of the path is exactly  $k$ .

We now show that  $\text{SUBSET-SUM} \leq_p \text{MP}$ . Given an instance  $(X, t)$  of SUBSET-SUM, we construct an instance  $(G, v, k)$  of MP as follows. The multigraph  $G$  contains a single vertex  $v$ . For each  $x_i \in X$ , we add a self-edge  $e_i$  from  $v$  to  $v$  with weight  $x_i$ . We set the target weight  $k$  for MP equal to  $t$ .

We now show that a yes instance of SUBSET-SUM exists iff a yes instance to MP exists.

$\Rightarrow$  Claim 1: if  $(X, t) \in \text{SUBSET-SUM}$ , then  $(G, v, k) \in \text{MP}$ .

If  $(X, t) \in \text{SUBSET-SUM}$ , then there exists a subset of elements  $X_0 = \{x_{i_1} \dots x_{i_k}\}$  whose sum is  $t$ . Consider the path in  $G$  that traverses edges  $e_{i_1} \dots e_{i_k}$ , each exactly once. This is a valid path that starts and ends at  $v$ , traverses no edge more than once, and, by our choice of edge weights, has total length exactly  $t$ . QED

$\Leftarrow$  Claim 2: if  $(G, v, k) \in \text{MP}$ , then  $(X, t) \in \text{SUBSET-SUM}$ .

Suppose there exists a suitable path  $p$  in  $G$  with length  $k$ . Let the sequence of edges on this path be  $\{e_1 \dots e_k\}$ . To each edge  $e_j$ , there corresponds an element  $x_j \in X$  such that  $w(e_j) = x_j$ . Let  $X_0$  be the set of all  $x_j$  corresponding to the edges of  $p$ . No edge is repeated, so the sum of all values in  $X_0$  is exactly  $k$ , which by construction is equal to  $t$ . Hence,  $(X, t) \in \text{SUBSET-SUM}$ .

QED