5.6 More NP-complete problems

3SAT

**Instance:** A CNF-formula $F$ such that each clause contains three literals.

**Question:** Is there a satisfying truth assignment for $F$?

**Theorem 5.10** 3SAT is NP-complete.

We know already that CNF-SAT belongs to NP, so 3SAT belongs to NP. Thus, to prove Theorem 5.10 it suffices to show that CNF-SAT $\leq_{p}^{m}$ 3SAT. Consider the mapping $g$ whose input is an arbitrary conjunction of clauses $F$ and whose output is given as follows: Replace each clause of $F$

$$\{x_1 \lor \ldots \lor x_n\}$$

with the following conjunction of clauses

$$\left(x_1 \lor x_2 \lor y_1\right) \land \left(x_3 \lor y_1 \lor y_2\right) \land \left(x_4 \lor y_2 \lor y_3\right) \land \ldots \land \left(x_{n-1} \lor x_n \lor y_{n-1}\right),$$

(5.7)

where $y_1, \ldots, y_{n-1}$ are new variables that do not occur in $\text{VAR}(F)$. We leave it to the reader to verify that the formula $F'$ is satisfiable if and only if the output formula $g(F)$ is satisfiable. The following observations will help in this task. Let $t$ be an arbitrary assignment to $\text{VAR}(F)$ and let $t'$ be any assignment to $\text{VAR}(F) \cup \{y_1, \ldots, y_{n-1}\}$ that agrees with $t$ on $\text{VAR}(F)$. (That is, if $u$ is a variable in $\text{VAR}(F)$, then $t(u) = t'(u)$.) Then, $t$ satisfies the clause in equation 5.7 if and only if $t'$ satisfies the formula in equation 5.6. Conversely, any assignment that satisfies equation 5.6 must also satisfy equation 5.7. Since $g$ is a polynomial-time reduction from CNF-SAT to 3SAT, we conclude that 3SAT is NP-complete.

**Example 5.2** ($k = 4$) Let $x_1, \ldots, x_4$ be variables and let $t$ be an assignment that assigns the value 1 to at least one of these variables, so that $t$ satisfies the clause

$$\left(x_1 \lor x_2 \lor x_3 \lor x_4\right).$$

Then, every extension of $t$ to the variables $\{x_1, \ldots, x_4, y_1\}$ satisfies the formula

$$\left(x_1 \lor x_2 \lor y_1\right) \land \left(x_3 \lor x_4 \lor y_3\right).$$

Conversely, every satisfying assignment to this formula must assign the value 1 to at least one of the $x_i$, $1 \leq i \leq 4$.

**Example 5.3** ($k = 5$) The same properties as in the previous example apply to the clause

$$\left(x_1 \lor x_2 \lor x_3 \lor x_4 \lor x_5\right)$$

and the corresponding conjunction of clauses

$$\left(x_1 \lor x_2 \lor y_1\right) \land \left(x_3 \lor y_1 \lor y_2\right) \land \left(x_4 \lor y_2 \lor y_3\right).$$

Some of the most famous NP-complete problems are about graphs. The following problem, VERTEX COVER, is NP-complete, and is an important tool for showing NP-completeness of other NP-complete problems. A vertex cover of a graph $G = (V,E)$ is a subset $V'$ of $V$ that, for each edge $(u,v) \in E$, contains at least
5.6 More NP-complete problems

![Graph](image.png)

**Figure 5.1** A graph $G$: $\{1,3,5\}$ and $\{1,2,4\}$ are vertex covers. Does $G$ have a vertex cover of size 2?

one of the adjacent vertices $u$ and $v$. The size of a vertex cover $V'$ is the number of distinct vertices it contains. These notions are illustrated in Figure 5.1.

**Vertex Cover**

Instance A graph $G = (V,E)$ and a positive integer $k \leq |V|$.

Question Is there a vertex cover of size $\leq k$ for $G$.

**Theorem 5.11** \textsc{Vertex Cover} is NP-complete.

**Proof** It is easy to see that \textsc{Vertex Cover} belongs to NP: Given a graph $G = (V,E)$, guess a set of vertices $V'$, and check that $V'$ is a vertex cover. This test can be performed deterministically in polynomial time.

Now we show that 3SAT $\leq^p_p$ \textsc{Vertex Cover}. We will describe a polynomial time-bounded construction that maps an instance $F$ of 3SAT to some graph $G = (V,E)$ and positive integer $k$ such that $G$ has a vertex cover of size $\leq k$ if and only if $F$ is satisfiable. The construction will consist of three components.

Let $U = \text{VAR}(F)$. For each variable $u_i \in U$, $V$ will contain vertices $u_i$ and $\overline{u_i}$, and $E$ will contain the edge $(u_i, \overline{u_i})$. This is called the \textit{truth setting component}. Note that any vertex cover will have to contain at least one of $u_i$ and $\overline{u_i}$.

Let $C$ be the set of clauses in $F$; that is, $F = \bigwedge_{i \in C} c_j$. For each clause $c_j \in C$, $V$ will contain three vertices $a_1[j]$, $a_2[j]$, and $a_3[j]$, and three edges joining them to make a triangle:

$$(a_1[j], a_2[j]), (a_2[j], a_3[j]), (a_3[j], a_1[j]).$$
this is called the satisfaction testing component. Note that any vertex cover will contain at least two vertices from each triangle.

The communications component adds edges between the satisfaction testing and truth setting components. This is the only component that depends on which literals are contained in which clauses. For each clause \( c_j \in C \), where \( c_j = (x_i \lor y_j \lor z_k) \), edges are added as follows:

\[(a_1[j], s_1), (a_2[j], y_j), (a_3[j], z_k)\].

This completes the definition of \( G \). The constant \( k \) is defined to be

\[ k = ||U|| + 2||C||. \]

Clearly, the construction takes polynomial time.

We need to show that \( G \) has a vertex cover of size \( \leq k \) if and only if \( F \) is satisfiable. Suppose \( V' \) is a vertex cover for \( G \) and \( ||V'|| \leq k \), then, as we have noted, \( V' \) has at least one vertex for each variable (i.e., at least one of \( u_i \) and \( \overline{u}_i \)), and has at least two vertices for each clause (i.e., at least two vertices from each triangle). Thus, \( ||V'|| = k \). Define an assignment \( t: U \rightarrow \{0, 1\} \) by \( t(u_i) = 1 \), if \( u_i \in V' \), and \( t(\overline{u}_i) = 0 \), if \( u_i \not\in V' \). We claim that this assignment satisfies each clause \( c_j \in C \).

Consider the triangle in the satisfaction testing component that corresponds to \( c_j \). Exactly two of the vertices of this triangle belong to \( V' \). The third vertex does not belong to \( V' \), so the communications component edge between this vertex and a vertex \( x \) in the truth setting component must be covered by the vertex in the truth setting component. By definition of the communications component, this means that \( c_j \) contains a literal \( x \in \{u_i, \overline{u}_i\} \), and that \( t(x) = 1 \). Thus, \( t \) satisfies \( c_j \).

Conversely, suppose that an assignment \( t \) satisfies each clause \( c_j \) in \( C \). For each variable \( u_i \in U \), either \( t(u_i) = 1 \) or \( t(\overline{u}_i) = 1 \). Place the vertex \( u_i \) into \( V' \), if \( t(u_i) = 1 \), and place the vertex \( \overline{u}_i \) into \( V' \), if \( t(\overline{u}_i) = 1 \). Then, \( V' \) contains one vertex of each edge in the truth setting component. In particular, if \( x \in \{u_i, \overline{u}_i\} \) is a literal in \( c_j \) that is assigned the value 1, then \( x \) is a vertex that is placed into \( V' \). By definition of the communications component, one vertex of the triangle in the satisfaction testing component that corresponds to \( c_j \) is covered by the edge that connects the triangle to the vertex \( x \). For each clause \( c_j \), place the other two vertices of the corresponding triangle into \( V' \). It follows that \( ||V'|| \leq k \), and that \( V' \) is a vertex cover. This completes the proof.

Figure 5.2 shows the graph that is obtained by applying the construction to the instance \( (u_1 \lor \overline{u}_2 \lor \overline{u}_3) \land (u_1 \lor u_2 \lor \overline{u}_3) \) of 3SAT, and shows the vertex cover that corresponds to the satisfying assignment \( t(u_1) = t(u_2) = t(u_3) = 1 \) and \( t(\overline{u}_3) = 0 \).

For any graph \( G \), recall from Example 1.4 that a clique is a complete subgraph of \( G \). Now that we know that VERTEX COVER is NP-complete, it is rather easy to show that the following CLIQUE problem is NP-complete.

**CLIQUE**
FIGURE 5.2. The instance of VERTEX COVER that results from the instance \((u_1 \lor \overline{u_2} \lor \overline{u_4}) \land (\overline{u_1} \lor u_2 \lor \overline{u_3})\) of 3SAT, with the vertex cover that corresponds to the assignment \(t(u_1) = t(u_2) = t(u_3) = 1\) and \(t(u_4) = 0\).
instance A graph \( G = (V, E) \) and a positive integer \( j \leq ||V|| \).
question Does \( G \) contain a clique of size \( j \) or more?

**Theorem 5.12** CLIQUE is NP-complete.

**Proof** It is easy to see that CLIQUE belongs to NP: To summarize the approach
given in Example 1.3, given a graph \( G \) and integer \( j \leq ||V|| \), guess a subgraph of \( G \) of size \( \leq j \), and then determine whether it is a clique.

Now we show that VERTEX COVER \( \leq^P \) CLIQUE. The complement of a graph \( G = (V, E) \) is the graph \( G^\prime = (V, E^\prime) \), where \( E^\prime = \{ (u, v) \mid u \in V, v \in V, \text{ and } (u, v) \notin E \} \). Given an instance of VERTEX COVER, a graph \( G \) and positive integer \( k \leq ||V^\prime|| \), the output of the polynomial time reduction is \( G^\prime \) and integer \( ||V^\prime|| - k \).

First we show that if \( V^\prime \) is a vertex cover for \( G \), then \( V^\prime - V \) is a clique of \( G^\prime \). Let \( V^\prime \) be a vertex cover for \( G \), and let \( u \) and \( v \) belong to \( V - V^\prime \). Since every edge of \( G \) has at least one adjacent vertex in \( V^\prime \), it follows that \((u, v) \notin E \). Thus, \((u, v) \in E^\prime \). and this proves the claim. It follows from the same line of reasoning that if \( V^\prime \) is a clique in \( G^\prime \), then \( V - V^\prime \) is a vertex cover for \( G \). 

**Homework 5.14** Show the natural reduction from CLIQUE to VERTEX COVER.

We conclude this chapter with mention of three additional NP-complete problems, 3-DIMENSIONAL MATCHING, PARTITION, and HAMILTONIAN CIRCUIT.

**3-DIMENSIONAL MATCHING**

instance A set \( M \subseteq W \times X \times Y \), where \( W, X, \) and \( Y \) are disjoint sets having the same number \( q \) of elements.
question Is there a subset \( M^\prime \) of \( M \), called a matching, such that \( ||M^\prime|| = q \) and no two elements of \( M^\prime \) agree in any coordinate?

**PARTITION**

instance A finite set \( A \) and a positive integer "size" \( s(a) \) for each \( a \in A \).
question Is there a subset \( A^\prime \) of \( A \) such that \( \sum_{a \in A} s(a) = \sum_{a \in A^\prime} s(a) \).

**HAMILTONIAN CIRCUIT**

instance A graph \( G = (V, E) \).
question Does \( G \) contain a Hamiltonian circuit?

The size of a proposed solution to the HAMILTONIAN CIRCUIT problem is within a polynomial of the size of an input graph, and verification of a proposed solution takes polynomial time. It follows that HAMILTONIAN CIRCUIT belongs to NP. The VERTEX COVER problem is used to show completeness of the HAMILTONIAN CIRCUIT problem, i.e., VERTEX COVER \( \leq^P \) HAMILTONIAN CIRCUIT.