*This document consists of thirty-three exercises on proving reductions and* NP*-completeness, generally ordered in increasing order of difficulty, but this is a matter of taste. The later exercises require more creativity than the earlier exercises. As such, I highly recommend you try and solve the earlier exercises before attempting the later exercises.*

*At the end of the document is a set of hints for each problem, including what problem(s) you should try and reduce from. In addition, extra exercises with fully worked solutions can be found at the end of the document. This is simply to show you the structure of a correct proof. No fully worked solutions will be provided for the problems in the problem sets.*

*Good luck, and may this serve to give you more confidence in the topic!*

## Reductions and Proving NP-completeness

1. Let $G = (V, E)$ be an undirected graph. A *Hamiltonian path* is a simple path that visits every vertex *exactly* once. It is known that deciding if a graph has a Hamiltonian path is NP-complete. Consider the related problem of deciding if there exist a simple path in $G$ that misses 2024 vertices; in other words, we want to decide if a graph $G$ has a simple path that visits $|V| - 2024$ vertices *exactly* once. Show that this problem is NP-complete by reducing from the Hamiltonian path problem.

   **Note.** *You may assume that an input graph has at least 2024 vertices; otherwise, return NO.*

2. Let $G = (V, E)$ be an undirected graph. Prove that deciding if $G$ has a simple path that passes through *at least* half of the vertices in $G$ is NP-complete.

3. We now generalise the previous problems. Instead of deciding if a graph contains a simple path that misses 2024 vertices, it now misses $k$ vertices, where $k$ is now part of the input. Prove that this problem is NP-complete.

   **Note.** *You can assume that $k$ is an integer that is polynomial in the size of the graph. This just removes some of the technical subtleties in the construction of the graph, which is not the point!*

4. Let $G = (V, E)$ be an undirected graph. Consider the following concepts on $G$.

   - An *independent set* $S$ of size $k$ is a subset of $k$ vertices such that no pair of vertices in $S$ are connected by a direct edge.
   - A *clique* $C$ of size $k$ is a subset of $k$ vertices such that every pair of vertices in $C$ are connected by a direct edge.

   It is known that deciding if a graph has an independent set of size at least $k$ is NP-complete. Use this to show that deciding if a graph has a clique of size at least $k$ is also NP-complete.

5. We reuse the concepts from the previous question. Let $G = (V, E)$ be an undirected graph. A *clique-independent* set $X$ of size $k$ is a subset of $2k$ vertices such that:

   - $X$ contains $k$ vertices that form an independent set $I$ of $G$,
   - $X$ contains $k$ vertices that form a clique $C$ of size $k$, and
   - $I \cap C = \emptyset$; that is, the independent set and clique are disjoint.

Prove that deciding if a graph has a clique-independent set of size at least $k$ is NP-complete.

6. Let $G = (V, E)$ be an undirected graph. Prove that deciding if there exist a simple path of length at least $k$ is NP-complete, where we measure the length of the path by the number of edges in the path.

7. Let $G = (V, E)$ be an undirected graph. A *high-degree independent set* of degree $k$ is an independent set $I$ where every vertex in $I$ has degree at least $k$ in $G$. Prove that deciding if a graph has a high-degree independent set of degree at least $k$ is NP-complete.

8. Let $G = (V, E)$ be an undirected graph. Prove that deciding if $G$ contains a spanning tree where every vertex in the spanning tree has degree at most 2024 is NP-complete. As an exercise, generalise this result!

9. Let $G = (V, E)$ be an undirected graph. A *proper 3-colouring* of $G$ is an assignment of vertices to *at most* three colours so that any pair of adjacent vertices do not share the same colour. It is known that deciding if a graph has a proper 3-colouring is NP-complete. Use this to show that deciding if a graph has a proper 4-colouring is NP-complete.

   **Note.** *A proper 4-colouring is an assignment of vertices to at most four colours with the same condition as a 3-colouring.*

10. Use the same idea as above to prove that deciding if a graph has a proper $k$-colouring is NP-complete for $k \geq 3$.

11. Let $A$ be an array of $n$ integers. We consider the following two problems on $A$.

    - SubsetSum: Given an array $A$ of $n$ integers and an integer $T$, decide if there exist a subset of integers in $A$ that sum to $T$.
    - Partition: Given an array $A$ of $n$ integers, decide if there exist a way to partition $A$ into two sets $E_1$ and $E_2$ such that

    $$\sum_{i \in E_1} A[i] = \sum_{i \in E_2} A[i].$$

    It is known that Partition is NP-complete. Use this to show that SubsetSum is also NP-complete. See Problem 33 for another proof of the NP-completeness of SubsetSum.

12. Consider the following problems.

    - HamiltonianCycle: Given an undirected and unweighted graph $G = (V, E)$, decide if there exist a simple cycle in $G$ that visits every vertex exactly once.
    - TravellingSalesperson: Given a complete and weighted graph $G$ and an integer $k$, decide if there exist a simple cycle in $G$ that visits every vertex exactly once with total weight at most $k$.

    It is known that HamiltonianCycle is NP-complete. Use this to show that TravellingSalesperson is also NP-complete.

    **Note.** *The input to the Travelling Salesperson problem is a complete graph, not an arbitrary graph.*

13. A *balloon subgraph* of size $k$ is an undirected graph that consists of a simple cycle of length $k$ and a simple path of size $k$, where the path contains an endpoint that lies on the cycle. Every other vertex on the cycle is disjoint from every other vertex on the path. A balloon of size $k$ has $2k$ vertices, $k$ vertices in the cycle and $k+1$ vertices in the path including the endpoint in the cycle.

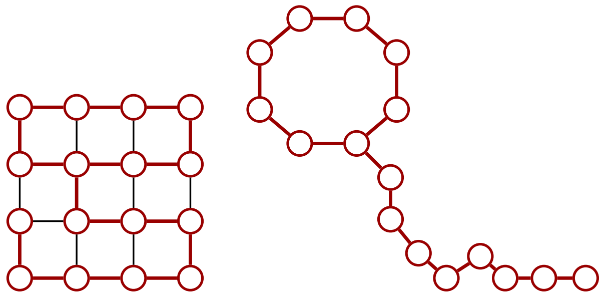


Figure 1: *A balloon subgraph of size* 8 *on the* $4 \times 4$ *grid graph.*

Given an undirected graph $G$ and an integer $k$, prove that it is NP-complete to decide if $G$ has a balloon subgraph of size at least $k$.

14. A *kite subgraph* of size $k$ is an undirected graph that consists of a clique of size $k$ and a simple path of size $k$, where the path contains an endpoint that lies on the clique. Every other vertex on the clique is disjoint from every other vertex on the path. A balloon of size $k$ has $2k$ vertices, $k$ vertices in the cycle and $k+1$ vertices in the path including the endpoint in the cycle.

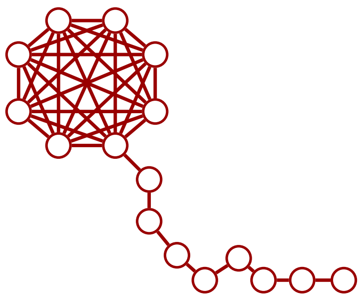


Figure 2: *A kite subgraph of size* 8.

Given an undirected graph $G$ and an integer $k$, prove that it is NP-complete to decide if $G$ has a kite subgraph of size at least $k$.

15. A *dandelion* of length $k$ is an undirected graph that consists of a simple path of length $k$, followed by $k$ additional vertices connected to one of the endpoints.

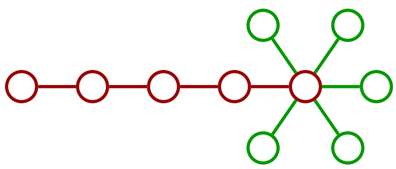


Figure 3: *A dandelion of length* 5.

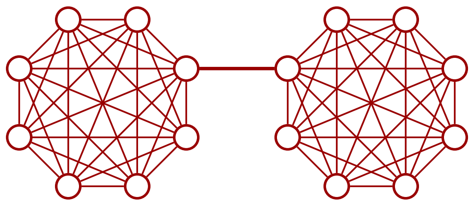
Given a graph $G$ and integer $k$, prove that it is NP-complete to decide if $G$ has a dandelion subgraph of size at least $k$.

16. Let $G = (V, E)$ be an undirected graph. A *bicycle covering* of $G$ is a pair of simple cycles $C_1, C_2$ in $G$ such that
    - $|C_1| = |C_2|$ and
    - every vertex in $G$ belongs to exactly one of the two cycles.

    In other words, the cycles partition $G$ and they have the same length. Prove that deciding if $G$ has a bicycle covering is NP-complete.

17. A *$k$-dumbbell* is a collection of two disjoint cliques $C_1$ and $C_2$, each of size $k$, together with an edge connecting $C_1$ and $C_2$.

    For example, the following is a dumbbell of size 8.

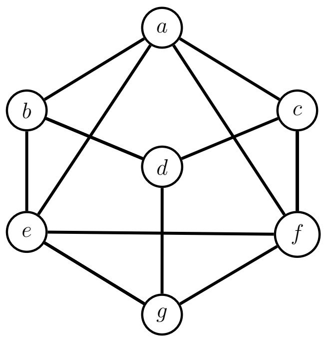


    We define the problem Dumbbell as follows: given an undirected graph $G = (V, E)$ and an integer $k$, decide whether there exist a $k$-dumbbell subgraph in $G$.

    Show that Dumbbell is NP-complete.

18. Let $G = (V, E)$ be an undirected graph. We consider a slightly weaker variant of the Hamiltonian path problem, where instead of finding a simple path visiting every vertex exactly once, we allow two vertices be visited exactly twice. In other words, we want to decide if there exist a walk in $G$ that visits two vertices exactly twice and every other vertex exactly once. Prove that this is still NP-complete.

19. A *doubly Hamiltonian tour* is a closed walk in an undirected graph $G$ that visits every vertex exactly twice. For example, the following graph

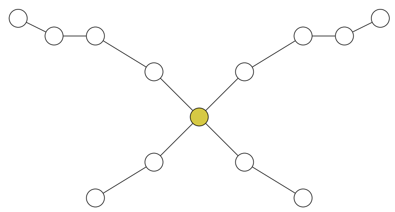



    has the doubly Hamiltonian tour

$$a \to b \to d \to g \to e \to b \to d \to c \to f \to a \to c \to f \to g \to e \to a.$$

    Given an undirected graph $G$, prove that deciding whether $G$ contains a doubly Hamiltonian tour is NP-complete.

20. Let $G = (V, E)$ be an undirected graph. A *spanning spider* $S = (V, E')$ is a spanning tree of $G$ such that there is at most one vertex with degree at least 2 in the induced subgraph $G[S]$; that is, there is at most one vertex of degree at least 2 in the graph generated by removing edges not in $E'$.



*A spanning spider where the vertex of degree 2 or more is highlighted.*

Prove that deciding if $G$ has a spanning spider is NP-complete.

21. We explore a slight variant of the Hamiltonian cycle problem. Here, we have a *weighted* graph. A *heavy-tonian cycle* in $G$ is a Hamiltonian cycle whose total weight of edges is at least half of the total weight of all edges in $G$. Prove that deciding whether $G$ contains a heavy-tonian cycle is NP-complete.

22. Let $G = (V, E)$ be an undirected graph. Prove that deciding if $G$ has a spanning tree with at most 2024 leaves is NP-complete.

23. Consider the following problem.

    HalfClique

    **Input.** An undirected graph $G = (V, E)$ on $2n$ vertices.

    **Task.** Decide if $G$ has a clique of size $n$.

    Show that HalfClique is NP-complete.

24. Consider the following problem.

    CliquePartition

    **Input.** An undirected graph $G = (V, E)$ and an integer $k$.

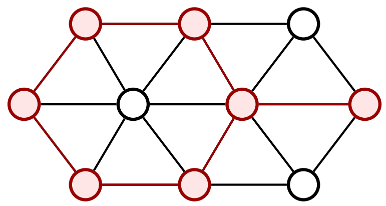
    **Task.** Decide if there exist a partition of $V$ into sets $V_1, \dots, V_k$ such that for each $1 \le i \le k$, $V_i$ forms a clique.

    Prove that CliquePartition is NP-complete.

25. Recall that in Question 11 we had a reduction from Partition to SubsetSum. Prove that Partition is NP-complete by reducing from SubsetSum.

    **Note.** *There is a reason why this is a lot later in the problem set! This direction is much harder. Here, we need to relax the set condition because we aren't guaranteed that the elements in the subset sum are distinct, so you may assume that a subset can also contain duplicate elements as well.*

26. Let $G = (V, E)$ be an undirected and unweighted graph. A subset $S \subseteq V$ of vertices is *□-free* if, for every four vertices $u, v, w, x \in S$, at least one of $(u, v), (v, w), (w, x), (u, x)$ does not appear in the edge set of $G$. In other words, $G$ induced by $S$ has no cycle of length 4. Prove that deciding if $G$ has a □-free subset of $k$ vertices is NP-complete.



*A □-free subset of 7 vertices with its induced edges. Note that this is **not** the largest subset.*

27. We consider the variant of the clique problem as follows:

    PartitionedClique

    **Input.** A graph $G = (V, E)$, where $V$ is partitioned into $k$ sets $V_1, \ldots, V_k$.

    **Task.** Decide if $G$ has a clique of size $k$, containing one vertex from each set $V_i$.

    Prove that PartitionedClique is NP-complete.

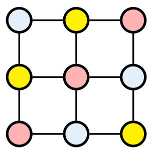
28. Let $G = (V, E)$ be an undirected graph. A *poor colouring* of $G$ assigns each vertex a set of two *distinct* colours such that each edge must use two different sets of colours. However, they may share a common colour. In other words, the endpoints of each edge in a poor colouring can share *at most* one common colour. Call a poor colouring that uses at most three colours a *poor 3-colouring*.

    Prove that deciding if $G$ has a poor 3-colouring is NP-complete.

29. Let $G = (V, E)$ be an undirected graph. A *rich colouring* assigns each vertex a set of two colours such that each edge must use two different sets of colours. Additionally, they may not share any colour. In other words, the endpoints of each edge in a rich colouring uses *exactly* four distinct colours. Therefore, every rich colouring is also a poor colouring. Call a rich colouring that uses at most five colours a *rich 5-colouring*.

    Prove that deciding if $G$ has a rich 5-colouring is NP-complete.

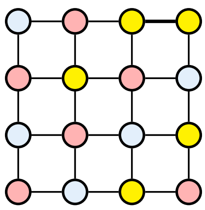
30. Let $G = (V, E)$ be an undirected graph. Recall that a proper 3-colouring is a colouring of the vertices of $G$ using at most three colours, so that no pair of adjacent vertices share the same colour. We explore a slightly weaker variant, which we call Balanced3Col. In this problem, we are still given an undirected graph $G$ but this time, we colour the vertices such that each colour has the same number of vertices. In a balanced 3-colouring, we still require any pair of adjacent vertices to not share the same colour.

    Prove that Balanced3Col is still NP-complete.

31. Let $G$ be an undirected graph. Recall that a proper 3-colouring is a colouring of the vertices of $G$ using at most three colours, so that no pair of adjacent vertices share the same colour.

    In this problem, we explore a slightly weaker variant, which we call Nearly3Col. In this problem, we are still given an undirected graph $G$ but this time, we allow *at most* one edge's endpoints in $G$ to have the same colour.

    Prove that Nearly3Col is still NP-complete.

32. Consider the following problem.

3DimensionalMatching

**Input.** Three sets $X, Y, Z$, each of size $n$, and a set $T \subset X \times Y \times Z$ of ordered triplets; that is, $T$ is composed of triplets $(x, y, z)$ such that $x \in X, y \in Y, z \in Z$.

**Task.** Decide if there exist a set $S \subseteq T$ of $n$ triplets in $T$ such that each element of $X, Y, Z$ is contained in *exactly* one triplet in $S$.

Prove that 3DimensionalMatching is NP-complete.

**Note.** *In some sense, the two-dimensional variant can be thought of as deciding if a bipartite graph has a perfect matching.*

33. Recall the SubsetSum problem from Question 11. Show that it is NP-complete by reducing from 3DimensionalMatching.

# Hints and Worked Examples

*There are no full solutions to the exercises; however, I have provided hints to all of the exercises. Attached is also a collection of worked examples to demonstrate the structure of a solution. If you spot any mistakes in the problem statement(s) or hints, please do not hesitate to email me at* `gerald.huang.1@unimelb.edu.au`. *Comments and feedback are appreciated!*

## Fully Worked Solutions

VERTEXCOVER

Let $G = (V, E)$ be an undirected graph. A *vertex cover* $C$ of size $k$ is a set of $k$ vertices such that every edge in $E$ is incident to at least one vertex in $C$. Prove that deciding if $G$ has a vertex cover of size at most $k$ is NP-complete.

**Hint.** *You may assume that the independent set problem is* NP*-complete.*

*Solution.* We prove that the problem is first in NP and then prove that the problem is NP-hard via a reduction from the independent set problem, which is known to be NP-complete.

- **Membership in NP.** Let $C$ be a set of $k$ vertices; it is easy to check the cardinality of $C$. Consider the following algorithm: for each edge $e \in E$, check that at least one of its endpoints is in $C$. If all of the edges have at least one endpoint in $C$, return YES.

  It is clear that this verifier runs in polynomial-time. Therefore, we have a polynomial-time verifier for the problem and so, VERTEXCOVER is in NP.

- **Membership in NP-hard.** We reduce from the independent set problem. Let $(G = (V, E), k)$ be an input to the independent set problem. Construct the input $(H = (V', E'), k')$ to the vertex cover problem, as follows:

  - Let $H = G$.
  - Set $k' = |V| - k$.

  It is firstly easy to see that this reduction can be done in polynomial-time. It now suffices to prove that $G$ has an independent set of size at least $k$ *if and only if* $H$ has a vertex cover of size at most $k' = |V| - k$.

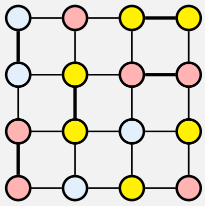
  - ( $\Longrightarrow$ ) Suppose that $G$ has an independent set $I$ of size at least $k$. We claim that $V \setminus I$ forms a vertex cover. Consider any edge $(u, v) \in E$ in $G$ for which $u, v \notin V \setminus I$. In particular, this implies that $u, v \in I$. But this implies that the independent set $I$ contains an edge between two vertices in $I$. This is a contradiction, which implies that at least one of $u$ or $v$ must be in $V \setminus I$. This implies that $V \setminus I$ forms a vertex cover of size at most $|V| - k = k'$.
  - ( $\Longleftarrow$ ) Suppose that $H$ has a vertex cover $C$ of size at most $k'$. Then every edge in $E'$ must be incident to at least one vertex in $C$. We claim that $I = V \setminus C$ forms an independent set. Suppose that there exist some edge $(u, v) \in E$ such that $u, v \in V \setminus C$. Then $C$ misses the edge $(u, v)$, which contradicts the assumption that $C$ is a vertex cover. Therefore, there cannot be an edge between two vertices in $V \setminus C$. This implies that $V \setminus C$ forms an independent set of size at least $|V| - k' = |V| - (|V| - k) = k$.

This implies that VertexCover is NP-hard. Together with the result that VertexCover is in NP, we conclude that VertexCover is NP-complete. □

**Almost3Col**

Let $G$ be an undirected graph. Recall that a proper 3-colouring is a colouring of the vertices of $G$ using at most three colours, so that no pair of adjacent vertices share the same colour.

In this problem, we explore a slightly weaker variant, which we call Almost3Col. In this problem, we are still given an undirected graph $G$ but this time, we allow each vertex $v$ to be adjacent to *at most* one other vertex of the same colour.



Prove that Almost3Col is still NP-complete.

*Solution.* We reduce from the 3-colouring problem by constructing a *clique* on six vertices as the gadget graph. We first prove membership in NP.

- **Membership in** NP: Given a colouring of $G$, we use the following verifier. For each vertex $v$, check that there is at most one vertex adjacent to $v$ that shares the same colour as $v$. We check through $\deg(v)$ neighbour vertices, and once we see that a vertex of the same colour has been seen, we just ensure that the remaining vertices have different colours. Once we check that every vertex has at most one neighbour vertex with the same colour, we just check that we have used at most 3 colours.

  There are at most $n - 1$ vertices in $G$ adjacent to $v$. Therefore, we have a linear-time algorithm to check each vertex and so, the verifier runs in $O(n^2)$ time which is polynomial-time.

- **Membership in** NP-hard: Let $G = (V, E)$ be an arbitrary graph; construct $G^* = (V^*, E^*)$ as follows:

  - For each vertex $v \in V$, construct five new vertices $v_1, v_2, v_3, v_4, v_5$ and connect them to $v$. Additionally, construct an edge between every pair of vertices to form a clique of six vertices, containing the vertices $v, v_1, v_2, v_3, v_4, v_5$.

  We claim that $G$ has a valid *3-colouring* if and only if $G^*$ has a valid *almost 3-colouring*.

  *Proof.* We prove both directions of the reduction.

  ($\implies$) Suppose that $G$ has a valid 3-colouring with colours red, blue, and yellow. We can extend this to a valid almost 3-colouring of $G^*$ as follows:

  – If $v$ is red, colour $v_1$ red, $v_2$ and $v_3$ yellow, and $v_4$ and $v_5$ blue.

  – If $v$ is yellow, colour $v_1$ yellow, $v_2$ and $v_3$ blue, and $v_4$ and $v_5$ red.

  – If $v$ is blue, colour $v_1$ blue, $v_2$ and $v_3$ red, and $v_4$ and $v_5$ yellow.

Since $G$ has a valid 3-colouring, no pair of adjacent vertices in $G$ share the same colour. Therefore, any vertex in $G$ is adjacent to at most one vertex (necessarily in the clique) with the same colour. Hence, $G^*$ has a valid almost 3-colouring.

( $\Longleftarrow$ ) Suppose that $G^*$ has a valid almost 3-colouring. For each vertex $v \in V$ of the original graph $G$, consider the gadget clique $v, v_1, v_2, v_3, v_4, v_5$. The clique must consist of exactly two vertices of each colour. Else, one colour must be assigned to at least three vertices which would break the almost 3-colouring of $G^*$. This implies that $v$ must be share the same colour as some vertex $v_i$ in the gadget graph. Since $G^*$ has a valid almost 3-colouring, any vertex $u$ adjacent to $v$ in $G$ must share different colours. In other words, the graph $G^*$, when restricted to only those vertices in $G$, must indeed have a valid 3-colouring.

$\square$

Since ALMOST3COL is in NP and NP-hard, we conclude that it is NP-complete. $\square$

## Hints

1. Try adding a few vertices in $G$ to form $G'$, so that if $G$ has a Hamiltonian path, then the same path must miss the vertices you added in $G'$. How many vertices should you add?
2. Imitate the previous argument, but you want to add some number of vertices in terms of the number of vertices in the input graph.
3. Generalise the above arguments.
4. Try proving that an independent set of $G$ corresponds to a clique of $\overline{G}$, where $\overline{G}$ is the complement graph. The converse is true as well!
5. The easiest reduction reduces from the clique problem. Try adding a few vertices to form an independent set.
6. A graph has a simple path of length $n - 1$ if and only if the graph has a Hamiltonian path.
7. Reduce from the independent set problem. Try blowing up the degree of every vertex in $G$. It might help to do this on paper for small $k$ first and then generalise the argument.
8. Reduce from the Hamiltonian path problem. Any spanning tree where every vertex has degree at most 2 in the tree is a Hamiltonian path! Try blowing up the degree of the vertices and repeat the same ideas as the previous problem.
9. If $G$ has a proper 3-colouring, add a few vertices to force any such colouring to use the additional colour.
10. Generalise the above argument.
11. Let $E_1$ and $E_2$ be a partition of $A$, so that the sums are the same. Choose a particular value for $T$ so that either $E_1$ or $E_2$ form a subset sum that sum to $T$.
12. It is easy to generate a complete graph from the input graph $G$. But you need to give weights to the edges of the complete graph appropriately, so that any Hamiltonian cycle has weight at most $k$, where you can freely choose $k$.

13. Reduce from Hamiltonian cycle.
14. Reduce from the clique problem; imitate the same argument as above.
15. Reduce from the Hamiltonian path problem; take inspiration from the earlier arguments.
16. Reduce from the Hamiltonian cycle problem; make two copies of the graph. When does a bicycle covering exist?
17. Reduce from the clique problem; make two copies of the graph. When does a dumbbell exist in the graph you constructed?
18. Reduce from the Hamiltonian path problem; take inspiration from the previous argument.
19. Reduce from the Hamiltonian cycle problem.
20. Reduce from the Hamiltonian path problem; make two copies of the graph. When does a spanning spider exist in the graph you constructed?
21. Reduce from the Hamiltonian cycle problem; give the edges of the input graph weights, so that any Hamiltonian cycle in $G$ satisfies the condition to also be heavy-tonian. There is a hack for this problem...
22. Reduce from the Hamiltonian path problem.
23. Reduce from the clique problem. If $G$ has a clique of size $k$, add $n$ vertices and find a way to extend this to a clique of size $n$.
24. You can choose the particular value for $k$. Consider the complement graph. Every clique of $G$ is an independent set of $\overline{G}$, and each independent set also forms a colour class in a proper colouring. Therefore, reduce from the 3-colouring problem. Choose a specific value for $k$.
25. Add a particular element $x$ into the array. To figure out what element to add, consider what happens when a partition exists. If $E_1, E_2$ form a partition of $A$ so that their sums are equal, we know that

    $$\sum_{i \in E_1} A[i] = \sum_{i \in E_2} A[i], \ \sum_{i \in E_1} A[i] + \sum_{i \in E_2} A[i] = \sum_{i=1}^{n} A[i].$$

    One of $E_1$ or $E_2$ must contain the element $x$. Find an expression for $x$.
26. Reduce from the independent set problem.
27. Reduce from the clique problem; a pair of vertices in a clique might belong to the same vertex set. Make a few copies of the same vertex, where the copies all belong to different vertex sets.
28. Reduce from the 3-colouring problem.
29. Reduce from the 3-colouring problem; throw away two of the colours by introducing a new vertex that is connected to every vertex. Subdivide each edge by a path of length 3.
30. Reduce from the 3-colouring problem; make three copies of the graph. Argue why the new graph has a balanced 3-colouring if and only if the original graph has a proper 3-colouring.
31. Construct a gadget to describe the property that in a nearly 3-colouring of $X$, exactly one edge of $X$ has endpoints that have the same colour.

32. Reduce from 3SAT; encode each variable $x_i$ as a gadget composed of a triangle. Think about how to construct clause gadgets and finally, how we connect clause and variable gadgets.
33. Associate each edge with a $3n$-digit number.