Chapter 3: Proving NP-completeness

Results

- Six Basic NP-Complete Problems
- Some Techniques for Proving NP-Completeness
- Some Suggested Exercises
Six Basic NP-Complete Problems

- 3-SATISFIABILITY (3SAT)
- 3-DIMENSIONAL MATCHING (3DM)
- VERTEX COVER (VC)
- CLIQUE
- CLIQUE
- HAMILTONIAN CIRCUIT (HC)
- PARTITION

First, a tour…
3-SATISFIABILITY (3-SAT)

INSTANCE: Collection $C = \{c_1, c_2, \ldots, c_m\}$ of clauses on a finite set $U$ of variables such that $\{c_j\} = 3$, for $1 \leq j \leq m$.

QUESTION: Is there a truth assignment for $U$ that satisfies all the clauses in $C$?
3-DIMENSIONAL MATCHING (3-DM)

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: Does $M$ contain a matching, that is, a subset $M' \subseteq M$ such that $|M'| = q$ and no two elements of $M'$ agree in any coordinate?

Example #1:

$q = 3$ (not part of input technically)

$W = \{a_1, a_2, a_3\}$

$X = \{b_1, b_2, b_3\}$

$Y = \{c_1, c_2, c_3\}$

$M = \{ (a_1, b_2, c_1), (a_2, b_1, c_3), (a_3, b_3, c_2), (a_1, b_1, c_3), (a_2, b_2, c_1) \}$

Yes! The first 3 form a matching.

Note that every element in each set will appear in exactly one triple in $M'$. 
Example #2:

$q = 3$ (not part of input technically)

$W = \{a_1, a_2, a_3\}$

$X = \{b_1, b_2, b_3\}$

$Y = \{c_1, c_2, c_3\}$

$M = \{(a_1, b_1, c_1), (a_1, b_1, c_2), (a_2, b_1, c_3), (a_3, b_1, c_1)\}$

No! In fact $b_2$ and $b_3$ do not appear in any triple.
Vertex Cover, Independent Set and Clique

VERTEX COVER (VC)

INSTANCE: A Graph $G = (V, E)$ and a positive integer $K \leq |V|$. 

QUESTION: Is there a vertex cover of size $K$ or less for $G$, that is, a subset $V' \subseteq V$ such that $|V'| \leq K$ and, for each edge $\{u,v\} \in E$, at least one of $u$ and $v$ belongs to $V'$?

$K = 3$

![Graph diagram](image)
CLIQUE
INSTANCE: A Graph $G = (V, E)$ and a positive integer $J \leq |V|$.
QUESTION: Does $G$ contain a clique of size $J$ or more?

$J = 3$

\[
\begin{array}{ccc}
\text{u} & \text{v} & \text{w} \\
\text{x} & \text{z} & \text{w} \\
\end{array}
\]
HAMILTONIAN CIRCUIT (HC)

INSTANCE: A Graph $G = (V, E)$.

QUESTION: Does $G$ have a Hamiltonian circuit, that is an ordering $< v_1, v_2, \ldots, v_n >$ of the vertices of $G$, where $n = |V|$, such that $\{v_n, v_1\} \in E$ and $\{v_i, v_{i+1}\} \in E$ for all $i, 1 \leq i < n$.

How is a Hamiltonian circuit different from a tour, as in TSP?

TRAVELING SALESMAN

INSTANCE: Set $C$ of $m$ cities, distance $d(c_i, c_j) \in \mathbb{Z}^+$ for each pair of cities $c_i, c_j \in C$ positive integer $B$.

QUESTION: Is there a tour of $C$ having length $B$ or less, i.e., a permutation $<c_{\pi(1)}, c_{\pi(2)}, \ldots, c_{\pi(m)}>\) of $C$ such that:

$$\sum_{i=1}^{m-1} d(c_{\pi(i)}, c_{\pi(i+1)}) + d(c_{\pi(m)}, c_{\pi(1)}) \leq B?$$
A restricted version of satisfiability in which all instances have exactly 3 literals per clause.

**3-SATISFIABILITY (3-SAT)**

INSTANCE: Collection $C = \{c_1, c_2, \ldots, c_m\}$ of clauses on a finite set $U$ of variables such that $\{c_j\} = 3$, for $1 \leq j \leq m$.

QUESTION: Is there a truth assignment for $U$ that satisfies all the clauses in $C$?

Contrasted with the more general SAT:

**SATISFIABILITY (SAT)**

INSTANCE: Collection $C$ of clauses on a finite set $U$ of variables.

QUESTION: Is there a truth assignment for $U$ that satisfies all the clauses in $C$?
Theorem 3.1: 3-SAT is NP-Complete.

1) $3$-SAT $\in$ NP
2) SAT $\propto$ 3-SAT

$I = U, C$
SAT instance

$A$

$I' = U', C'$
3-SAT instance

$I$ is satisfiable iff $I'$ is
3-Satisfiability

- Each clause $c_j$ in $I \Rightarrow$ a set $C_j'$ of clauses with new variables $U_j'$

- Let $c_j = \{z_1, z_2, \ldots, z_k\}$ where $z_i$'s are literals derived from variables in $U$

- Case 1) $k=1$ \quad $c_j = \{z_1\}$
  \[
  U_j' = \{y_j^1, y_j^2\} \\
  C_j' = \{ \{z_1, y_j^1, y_j^2\}, \{z_1, \overline{y_j}^1, y_j^2\}, \{z_1, y_j^1, \overline{y_j}^2\}, \{z_1, \overline{y_j}^1, \overline{y_j}^2\} \}
  \]

- Case 2) $k=2$ \quad $c_j = \{z_1, z_2\}$
  \[
  U_j' = \{y_j^1\} \\
  C_j' = \{ \{z_1, z_2, y_j^1\}, \{z_1, z_2, \overline{y_j}^1\} \}
  \]
3-Satisfiability

- Case 3) $k=3$ \( c_j = \{z_1, z_2, z_3\} \)

  \[
  U_j' = \{\} \\
  C_j' = \{ \{z_1, z_2, z_3\} \}
  \]

- Case 4) $k\geq4$ \( c_j = \{z_1, z_2, \ldots, z_k\} \)

  \[
  U_j' = \{y_j^i \mid 1 \leq i \leq k-3\} \\
  C_j' = \{ \{z_1, z_2, y_j^1\} \cup \{\overline{y_j^i}, z_{i+2}, y_j^{i+1}\} \mid 1 \leq i \leq k-4\} \cup \overline{\{y_j^{k-3}, z_{k-1}, z_k\}}
  \]

Example:

\[
\begin{align*}
  c_j &= \{z_1, z_2, z_3, z_4, z_5, z_6\} \\
  U_j' &= \{ \{z_1, z_2, a\}, \{\overline{a}, z_3, b\}, \{b, z_4, c\}, \{c, z_5, z_6\} \}
\end{align*}
\]
Finally, let:

\[ U' = U \cup \{ U'_j | 1 \leq j \leq m \} \]
\[ C' = \{ C'_j | 1 \leq j \leq m \} \]

Claim:

1) \( I = U, C \) is satisfiable iff \( I' = U', C' \) is satisfiable.
2) \( I' \) can be computed in polynomial time.
A review of basic logic…and language…

*A if and only if B*

- Equates to:
  - *A if B*  
    - The "if" part
  - *A only if B*  
    - The "only if" part

- Similarly:

  - *A if B*  
    - ≈  
    - *B only if A*  
    - ≈  
    - \( B \Rightarrow A \)
  - *A only if B*  
    - ≈  
    - *B if A*  
    - ≈  
    - \( A \Rightarrow B \)
Six Basic NP-Complete Problems

- Lastly, to prove:
  
  \[ X \Rightarrow Y \]

- We frequently show:
  
  \[ \neg Y \Rightarrow \neg X \]
So let's go back to the basic claim:

1) \( I = U,C \) is satisfiable iff \( I' = U',C' \) is satisfiable.

   a) (if) \( I = U,C \) is satisfiable if \( I' = U',C' \) is…

   b) (only if) \( I = U,C \) is satisfiable only if \( I' = U',C' \) is…
2) $I'$ can be computed in polynomial time:

Let $m = |C|$, $n = |U|$

- case 1) creates 4 clauses
- case 2) creates 2 clauses
- case 3) creates 1 clause
- case 4) creates $k-1$ clauses, where $k$ is the number of literals in the clause

This gives a total of at most $k-1$ new clauses in $I'$ for each clause in $I$

Therefore, there is a total of $m*(k-1)$ clauses in $I'$

Since $k \leq 2n$, it follows that there are at most $2nm$ clauses in $I'$, which is $O(mn)$
3-SAT is an example of what is called a “special case” of SAT.

- Some special cases, like 3-SAT, are NP-complete
- Others are solvable in polynomial time (chapter 4)

How about 4-SAT?

- 5-SAT?
- N-SAT, for any fixed N>=3?
- 1-SAT?
- 2-SAT?
Not All Equal (NAE) 3-SAT

INSTANCE: Collection $C = \{c_1, c_2, \ldots, c_m\}$ of clauses on a finite set $U$ of variables such that $\{c_j\} = 3$, for $1 \leq j \leq m$.

QUESTION: Is there a truth assignment for $U$ such that each clause in $C$ has at least one true literal and at least one false literal?

Fact: NAE 3-SAT is NP-complete
Examples:

\[
\begin{align*}
\{ (a, b, c), (a, b, c), (a, b, c) \} & \quad a = T, b = T, c = F \\
\{ (a, b, c), (a, b, c), (a, b, c), (a, b, c), (a, b, c) \} & \quad \text{Not even satisfiable} \\
\{ (a, b, c), (a, b, c), (a, b, c), \bar{a}, \bar{b}, \bar{c} \} & \quad \text{Satifiable, but not with} \\
\{ (a, b, c), (a, b, c), (a, b, c), (a, b, c), (a, b, c) \} & \quad \text{at least one literal true and false per clause}
\end{align*}
\]
Hypergraph 2-Colorability (H2C)

INSTANCE: Hypergraph \( H = (V, E) \), where \( 2 \leq |e_i| \leq 3 \), for all \( e_i \in E \).

QUESTION: Is \( H \) 2-colorable? In other words, is there a function \( f : V \rightarrow \{0, 1\} \) such that for all \( e_i \in E \) there exist vertices \( u, v \in e_i \) such that \( f(u) \neq f(v) \)?

Examples:

No

Yes

No
Theorem: H2C is NP-complete

H2C ∈ NP

Not All Equal 3-SAT ∝ H2C
Suppose \( I = U, C \) where \( U = \{u_0, u_1, \ldots, u_{n-1}\} \) and \( C = \{c_0, c_1, \ldots, c_{m-1}\} \)

Construct \( H = (V, E) \) where:

\[
V = \{ v_i^1 \mid u_i \in U \} \cup \{ v_i^2 \mid u_i \in U \}
\]

- \( v_i^1 \) corresponds to \( u_i \)
- \( v_i^2 \) corresponds to the complement of \( u_i \)

\[
E = E_1 \cup E_2 \text{ where}
\]

\[
E_1 = \{ (v_1, v_2, v_3) \mid c_i \in C, c_i = (z_1, z_2, z_3) \text{ and } v_1, v_2, v_3 \text{ are the vertices corresponding to the literals } z_1, z_2 \text{ and } z_3, \text{ respectively} \}
\]

\[
E_2 = \{ (v_i^1, v_i^2) \mid 0 \leq i \leq n-1 \}
\]
Example:

\[ U = \{ u_0, u_1, u_2, u_3 \} \]
\[ C = \{ (u_0, u_1, u_2), (\overline{u_0}, \overline{u_1}, u_3), (u_1, u_2, \overline{u_3}) \} \]

The resulting hyper-graph is 2-colorable IFF there is a NAE satisfying truth assignment for the Boolean expression.

- (if) Suppose there is a NAE assignment, then there is a 2-coloring.
- (only if) Suppose there is a 2-coloring, then there is a NAE satisfying assignment.
3-DIMENSIONAL MATCHING (3-DM)

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: Does $M$ contain a matching, that is, a subset $M' \subseteq M$ such at $|M'| = q$ and no two elements of $M'$ agree in any coordinate?

Example #1:

$q = 3$ (not part of input technically)

$W = \{a_1, a_2, a_3\}$

$X = \{b_1, b_2, b_3\}$

$Y = \{c_1, c_2, c_3\}$

$M = \{(a_1, b_2, c_1), (a_2, b_1, c_3), (a_3, b_3, c_2), (a_1, b_1, c_3), (a_2, b_2, c_1)\}$

Yes! The first 3 form a matching.

Note that every element in each set will appear in exactly one triple in $M'$. 
Example #2:

q = 3 (not part of input technically)
W = \{a_1, a_2, a_3\}
X = \{b_1, b_2, b_3\}
Y = \{c_1, c_2, c_3\}

M = \{(a_1, b_1, c_1), (a_1, b_1, c_2), (a_2, b_1, c_3), (a_3, b_1, c_1)\}

No! In fact b_2 and b_3 do not appear in any triple.
**Theorem:** 3DM is NP-complete.

**Proof:**

1) $3DM \in NP$

2) $3$-SAT $\preceq 3DM$
Suppose \( I = U, C \) where:

\[
U = \{u_1, u_2, \ldots, u_n\} \\
C = \{c_1, c_2, \ldots, c_m\}
\]

Construct \( M \) as follows.

\( M \) will consists of three types of “components:”

- Truth Setting and fan-out
- Satisfaction testing
- Garbage collection
Suppose $I = U, C$ where:

$$U = \{u_1, u_2, \ldots, u_n\}$$
$$C = \{c_1, c_2, \ldots, c_m\}$$

Construct $M$ as follows.

$M$ will consists of three types of “components:”

- Truth Setting and fan-out
- Satisfaction testing
- Garbage collection
For each variable $u_i \in U$, create the following (TS&FO) elements:

$$T_i^t = \{ (\overline{u}_i[j], a_i[j], b_i[j]) \mid 1 \leq j \leq m \} \cup \{ (u_i[m], a_i[1], b_i[m]) \}$$

$$T_i^f = \{ (u_i[j], a_i[j+1], b_i[j]) \mid 1 \leq j \leq m-1 \} \cup \{ (u_i[m], a_i[1], b_i[m]) \}$$

a big “blob” for each Boolean variable (suppose $m=4$):
Truth Setting and Fan Out Components

- Observation #1: Creates 2mn elements in W, mn in X and mn in Y.
  - $u_i[j]$ represents the fact that variable $u_i$ could occur in clause $c_j$.
  - Similarly for $\bar{u}_i[j]$

- Observation #2: $a_i[j]$ and $b_i[j]$ both occur in exactly two triples and nowhere else.

- Observation #3: This tells us that in any matching $M' \subseteq M$ we must have all white or all shaded triples, in other words, all triples from $T_i^1$ or all triples from $T_i^\dagger$
  - This corresponds to setting the variable $false$ or $true$, respectively.
For each clause $c_j \in C$, create (ST) triples:

$$C_j = \{ (u_i[j], s_1[j], s_2[j]) | u_i \in c_j \} \cup \{ (\overline{u_i}[j], s_1[j], s_2[j]) | \overline{u_i} \in c_j \}$$

Observation #4: This adds $m$ elements to $X$, and $m$ elements to $Y$.

Observation #5: For each $j$, $1 \leq j \leq m$, $s_1[j]$ and $s_2[j]$ both appear in exactly three triples and nowhere else.

Observation #6: Any matching must choose exactly one triple from $C_j$
  - This corresponds to making that literal true, thereby satisfying the clause.
  - This gives a total of $m$ triples selected from the ST triples.

Observation #7: Any matching will contain exactly one triple from $C_j$ if and only if there is a satisfying truth assignment for the 3-SAT instance.
Garbage Collection Components

- For each clause $c_j \in C$, create $2mn(n-1)$ (GC) triples:

  $$G = \{ (u_i[j], g_1[k], g_2[k]), (\overline{u}_i[j], g_1[k], g_2[k]) \mid 1 \leq k \leq m(n-1), 1 \leq i \leq n, 1 \leq j \leq m\}$$

- Observation #8: This adds $m(n-1)$ elements to $X$, and to $Y$.

- Observation #9: For each $k$, $1 \leq k \leq m(n-1)$, $g_1[k]$ and $g_2[k]$ both appear in exactly $2mn$ triples and nowhere else.

- Observation #10: Consequently, exactly $m(n-1)$ triples from $G$ must occur in any matching for $M$.

- Observation #11: $|W| = |X| = |Y| = 2mn = q$. 


Garbage Collection Components

- The resulting instance I’ of 3DM has a matching IFF the 3-SAT instance I is satisfiable.

- (if) Suppose you have a satisfying truth assignment for I. How do you construct the matching?

- (only if) Suppose you have a matching. What is the satisfying truth assignment?
**VERTEX COVER (VC)**

INSTANCE: A Graph $G = (V, E)$ and a positive integer $K \leq |V|$.

QUESTION: Is there a vertex cover of size $K$ or less for $G$, that is, a subset $V' \subseteq V$ such that $|V'| \leq K$ and, for each edge $\{u,v\} \in E$, at least one of $u$ and $v$ belongs to $V'$?

$K = 3$
INDEPENDENT SET (IS)

INSTANCE: Graph $G = (V, E)$ and positive integer $J \leq |V|$.
QUESTION: Does $G$ contain an independent set of size $J$ or more?

$J = 2$
Vertex Cover, Independent Set and Clique

CLIQUE
INSTANCE: A Graph $G = (V, E)$ and a positive integer $J \leq |V|$.
QUESTION: Does $G$ contain a clique of size $J$ or more?

$J = 3$
Lemma 3.1: For any graph $G=(V,E)$ and subset $V' \subseteq V$, the following statements are equivalent:

(a) $V'$ is a vertex cover for $G$
(b) $V-V'$ is an independent set for $G$
(c) $V-V'$ is a clique in the complement $G^c$ of $G$, where $G^c=(V,E^c)$ with $E^c=\{ \{u,v\} \mid u,v \in V \text{ and } \{u,v\} \notin E \}$
Lemma 3.1: For any graph $G=(V,E)$ and subset $V' \subseteq V$, the following statements are equivalent:

(a) $V'$ is a vertex cover for $G$
(b) $V-V'$ is an independent set for $G$
(c) $V-V'$ is a clique in the complement $G^c$ of $G$, where $G^c=(V,E^c)$ with $E^c=\{\{u,v\} \mid u,v \in V \text{ and } \{u,v\} \notin E\}$
**Theorem:** VERTEX COVER is NP-complete

**Proof:**

1) VC ∈ NP

2) 3SAT ∝ VC

Suppose I = U, C is an instance of 3SAT where:

\[ U = \{u_1, u_2, \ldots, u_n\} \]
\[ C = \{c_1, c_2, \ldots, c_m\} \]

We will construct a graph G=(V,E) and a positive integer K≤|V| such that G has a vertex cover of size K or less if and only if C is satisfiable.
Vertex Cover

For each variable $u_i \in U$, there is a “truth-setting component” $T_i = (V_i, E_i)$ where:

$$V_i = \{u_i, \overline{u_i}\} \text{ and } E_i = \{\{u_i, \overline{u_i}\}\}$$

*Note that any vertex cover will have to contain at least one of $u_i$ and its complement.

For each clause $c_j \in C$, create a “satisfaction testing component” $S_j = (V_j', E_j')$:

$$V_j' = \{a_1[j], a_2[j], a_3[j]\}$$
$$E_j' = \{\{a_1[j], a_2[j]\}, \{a_1[j], a_3[j]\}, \{a_2[j], a_3[j]\}\}$$

*Note that any vertex cover will have to contain at least two vertices from $V_j'$ in order to cover the edges in $E_j'$. 
For each clause $c_j \in C$, let the three literals in $c_j$ be $x_j$, $y_j$, $z_j$. Then add “communication edges:”

$$E_j = \{ \{a_1[j], x_j\}, \{a_2[j], y_j\}, \{a_3[j], z_j\}\}$$

Finally, let $K = n + 2m$. 
Example:

\[ U = \{ u_1, u_2, u_3, u_4 \} \]
\[ C = \{ (u_1, \bar{u}_3, \bar{u}_4), (\bar{u}_1, u_2, \bar{u}_4) \} \]
\[ K = n + 2m = 8 \]
Observations:

- The transformation can be performed in polynomial-time.
- \( G = (V, E) \) will have a vertex cover of size \( K \) or less IFF \( I = U, C \) is satisfiable.

(if) Suppose \( I = U, C \) is satisfiable…

(only if) Suppose \( G = (V, E) \) has a vertex cover of size \( K \) or less…
Recall Lemma 3.1:

**Lemma 3.1:** For any graph $G=(V,E)$ and subset $V' \subseteq V$, the following statements are equivalent:

(a) $V'$ is a vertex cover for $G$
(b) $V-V'$ is an independent set for $G$
(c) $V-V'$ is a clique in the complement $G^c$ of $G$, where $G^c=(V,E^c)$ with $E^c=\{\{u,v\} \mid u,v \in V$ and $\{u,v\} \not\in E\}$

What does the previous result say about the independent set and clique problems?
INDEPENDENT SET (IS)
INSTANCE: Graph $G = (V, E)$ and positive integer $J \leq |V|$.
QUESTION: Does $G$ contain an independent set of size $J$ or more?

Prove IS NP-complete by giving a transformation from 3DM.

4-DIMENSIONAL MATCHING (4-DM)
INSTANCE: A set $M \subseteq W \times X \times Y \times Z$, where $W$, $X$, $Y$ and $Z$ are disjoint sets having the same number $q$ of elements.
QUESTION: Does $M$ contain a matching, that is, a subset $M' \subseteq M$ such that $|M'| = q$ and no two elements of $M'$ agree in any coordinate?

Prove 4-DM NP-complete by giving a transformation from 3DM.
HAMILTONIAN CIRCUIT (HC)

INSTANCE: A Graph $G = (V, E)$.

QUESTION: Does $G$ have a Hamiltonian circuit, that is an ordering $\langle v_1, v_2, \ldots, v_n \rangle$ of the vertices of $G$, where $n = |V|$, such that $\{v_n, v_1\} \in E$ and $\{v_i, v_{i+1}\} \in E$ for all $i$, $1 \leq i < n$.

How is a Hamiltonian circuit different from a tour, as in TSP?
HAMILTONIAN CIRCUIT (HC)
INSTANCE: A Graph $G = (V, E)$.
QUESTION: Does $G$ have a Hamiltonian circuit, that is an ordering $< v_1, v_2, \ldots, v_n >$ of the vertices of $G$, where $n = |V|$, such that $\{v_n, v_1\} \in E$ and $\{v_i, v_{i+1}\} \in E$ for all $i$, $1 \leq i < n$.

VERTEX COVER (VC)
INSTANCE: A Graph $G = (V, E)$ and a positive integer $K \leq |V|$.
QUESTION: Is there a vertex cover of size $K$ or less for $G$, that is, a subset $V' \subseteq V$ such that $|V'| \leq K$ and, for each edge $\{u, v\} \in E$, at least one of $u$ and $v$ belongs to $V'$?
Theorem: HC is NP-complete

Proof:
1) HC $\in$ NP
2) VC $\preceq$ HC

Let $G = (V,E)$ and $K \leq |V|$ be an instance of VC.

We will construct a graph $G' = (V',E')$ such that $G$ has a vertex cover of size $K$ or less if and only if $G'$ has a Hamiltonian circuit.
1) Add “selector” vertices \( a_1, a_2, \ldots, a_k \) to \( V' \)

2) For each edge \( e = \{u, v\} \) in \( E \), construct the following “cover testing” component:

More specifically, add the following 12 vertices:

\[
V'_e = \{(u,e,i), (v,e,i) : 1 \leq i \leq 6\}
\]

And 14 edges:

\[
E'_e = \{((u,e,i), (u,e,i+1)), ((v,e,i), (v,e,i+1)) : 1 \leq i \leq 5 \}
\]

\[
\cup \{((u,e,3), (v,e,1)), ((v,e,3), (u,e,1))\}
\]

\[
\cup \{((u,e,6), (v,e,4)), ((v,e,6), (u,e,4))\}
\]
3) For each vertex \( v \in V \), let the edges incident on \( v \) be:

\[
e_{v[1]}, e_{v[2]}, \ldots, e_{v[\text{deg}(v)]}
\]

Add the following edges:

\[
E'_v = \{(v, e_{v[i]}, 6), (v, e_{v[i+1]}, 1) : 1 \leq i < \text{deg}(v)\}
\]

This creates a path in \( G' \) that “touches” the cover-testing components corresponding to the edges adjacent to \( v \) in \( G \).

This path corresponds to the vertex \( v \) from \( G \).
4) Finally, add edges to $G'$ that connect the first ($s_i$) and last ($e_i$) vertices from each of these paths to every one of the selector vertices:

$$E'' = \{ \{a_i, (v, e_{v[1]}, 1)\}, \{a_i, (v, e_{v[\text{deg}(v)]}, 6)\} : 1 \leq i \leq K, v \in V \}$$
Hamiltonian Circuit (HC)

A complete example – consider the following instance of vertex cover:

\[ K = 3 \]
A complete example – consider the following instance of vertex cover:

\[ K = 3 \]
Hamiltonian Circuit (HC)

A complete example – consider the following instance of vertex cover:

\[ K = 3 \]
A complete example – consider the following instance of vertex cover:

\[ K = 3 \]
Hamiltonian Circuit (HC)

A complete example – consider the following instance of vertex cover:

\( K = 3 \)
A complete example – consider the following instance of vertex cover:

\[ K = 3 \]
A complete example – consider the following instance of vertex cover:

\[ K = 3 \]
Hamiltonian Circuit (HC)

A complete example – consider the following instance of vertex cover:

\[ K = 3 \]
A complete example – consider the following instance of vertex cover:

\[ K = 3 \]
A complete example – consider the following instance of vertex cover:

K = 3
Hamiltonian Circuit (HC)

A complete example – consider the following instance of vertex cover:

\[ K = 3 \]
A complete example – consider the following instance of vertex cover:

\[ K = 3 \]

\[ G \text{ as a vertex cover of size 3 if and only if } G' \text{ has a Hamiltonian circuit} \]
Hamiltonian Circuit (HC)

(only if)

**Figure 3.5** The three possible configurations of a Hamiltonian circuit within the cover-testing component for edge $e = (u, v)$, corresponding to the cases in which (a) $u$ belongs to the cover but $v$ does not, (b) both $u$ and $v$ belong to the cover, and (c) $v$ belongs to the cover but $u$ does not.
Hamiltonian Circuit (HC)

\[ K = 3 \]

Figure 3.5 The three possible configurations of a Hamiltonian circuit within the cover-testing component for edge \( e = \{u, v\} \), corresponding to the cases in which (a) \( u \) belongs to the cover but \( v \) does not, (b) both \( u \) and \( v \) belong to the cover, and (c) \( v \) belongs to the cover but \( u \) does not.
HAMILTONIAN PATH (HP)

INSTANCE: A Graph $G = (V, E)$, and vertices $u, v \in V$.

QUESTION: Is there a Hamiltonian path from vertex $u$ to vertex $v$, that is an ordering $v_1, v_2, \ldots, v_n$ of the vertices of $G$, where $n = |V|$, such that $u = v_1$, $v = v_n$, and $\{v_i, v_{i+1}\} \in E$ for all $i$, $1 \leq i < n$.

FACT: HP is NP-complete (see page 60).
INDEPENDENT SET (IS)

INSTANCE: Graph $G = (V, E)$ and positive integer $J \leq |V|$.

QUESTION: Does $G$ contain an independent set of size $J$ or more?

**Theorem:** IS is NP-complete

1) $IS \in NP$

2) $3DM \asymp IS$
Proof:

Let M be an instance of 3DM, and construct an instance of IS as follows.

\[ V = \{v_i \mid t_i \in M\} \]
\[ E = \{v_i, v_j \mid i \neq j \text{ and } t_i \text{ and } t_j \text{ agree on some coordinate}\} \]
\[ j = q \]

Claims:
1) The transformation can be performed in polynomial-time
2) M contains a matching if and only if \( G=(V,E) \) contains an independent set of size \( j \).
(only if)

Let $M' = \{t_0, t_1, \ldots, t_{q-1}\}$ be a matching, where $M' \subseteq M$. It can be easily verified that $V' = \{v_0, v_1, \ldots, v_{q-1}\}$ is an independent set of size $j$.

(if)

Let $V' = \{v_0, v_1, \ldots, v_{q-1}\}$ is an independent set of size $j$ for $G$. Then it can be easily verified that $M' = \{t_0, t_1, \ldots, t_{q-1}\}$ is a matching.
4-DIMENSIONAL MATCHING (4-DM)

INSTANCE: A set $M \subseteq W \times X \times Y \times Z$, where $W$, $X$, $Y$ and $Z$ are disjoint sets having the same number $q$ of elements.

QUESTION: Does $M$ contain a matching, that is, a subset $M' \subseteq M$ such at $|M'| = q$ and no two elements of $M'$ agree in any coordinate?

**Theorem:** 4DM is NP-complete

1) $4DM \in NP$

2) $3DM \propto 4DM$
Proof:

Let M, W, X, and Y be an instance of 3DM, and construct an instance of 4DM as follows:

\[
\begin{align*}
W' &= W \\
X' &= X \\
Y' &= Y \\
Z &= \{z_i, \text{ where } 0 \leq i \leq q-1, \text{ and } q = |W|=|X|=|Y|\} \\
M' &= \{(w_i,x_j,y_k,z_r) \mid (w_i,x_j,y_k) \in M \text{ and } z_r \in Z\}
\end{align*}
\]

In other words, for each triple in M and value in Z, add a quadruple in M’.

Claims:
1) The transformation can be performed in polynomial-time
2) M contains a matching if and only if M’ does.
(only if)
Let $M'' = \{t_0, t_1, \ldots, t_{q-1}\}$ be a 3D matching, where $M'' \subseteq M$.
Create a 4D matching $M'''$ by adding $z_i$ to triple $t_i$, for all $i$, where $0 \leq i \leq q-1$.
It can easily be verified that $M'''$ is a 4D matching for $I'$.

(if)
Let $M'' = \{t_0, t_1, \ldots, t_{q-1}\}$ be a 4D matching, where $M'' \subseteq M'$.
Create a 3D matching by letting $M''' = \{(w_i, x_j, y_k) \mid (w_i, x_j, y_k, z_r) \in M''\}$.
It can easily be verified that $M'''$ is a 3D matching for $I'$. 
PARTITION

INSTANCE: A finite set $A$ and a “size” $s(a) \in \mathbb{Z}^*$ for each $a \in A$.

QUESTION: Is there a subset $A' \subseteq A$ such that:

$$\sum_{a \in A'} s(a) = \sum_{a \in A - A'} s(a)$$

3-DIMENSIONAL MATCHING (3-DM)

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: Does $M$ contain a matching, that is, a subset $M' \subseteq M$ such at $|M'| = q$ and no two elements of $M'$ agree in any coordinate?
**Theorem:** Partition is NP-complete

**Proof:**

1) \( \text{PARTITION} \in \text{NP} \)

2) \( 3\text{DM} \preceq \text{PARTITION} \)

Let \( W, X, \) and \( Y \) with \( |W| = |X| = |Y| = q \), and \( M \subseteq W \times X \times Y \) be an instance of 3DM.

Also let:

\[
W = \{w_1, w_2, \ldots, w_q\} \\
X = \{x_1, x_2, \ldots, x_q\} \\
Y = \{y_1, y_2, \ldots, y_q\}
\]

and:

\[
M = \{m_1, m_2, \ldots, m_k\}
\]

where \( k = |M| \)
We will construct a set $A$, and a size $s(a) \in \mathbb{Z}^+$ for each $a \in A$.

This set $A$ will be such that $A$ contains a subset $A'$ satisfying:

$$\sum_{a \in A'} s(a) = \sum_{a \in A - A'} s(a)$$

if and only if $M$ contains a matching.
First add \( k \) elements \( \{a_i : 1 \leq i \leq k\} \) to \( A \), where \( a_i \) corresponds with triple \( m_i \in M \), and the size \( s(a_i) \) in binary is given by:

\[
p = \left\lceil \log_2 \left( \log_2 \left( k + 1 \right) \right) \right\rceil
\]

Notes:

- A triple from \( M \) corresponds to a 3-bit binary number, where each bit is in the rightmost position of one sub-field.
- If all the binary numbers were added up for all the triples, no sub-field would overflow (because of the choice of \( p \)).
Let:

\[ B = \sum_{i=0}^{3q-1} 2^p_j \]

- Then \( B \) corresponds to the binary number containing a 1 in every sub-field rightmost position.

- If a set \( M' \) of triples formed a matching, then the corresponding binary numbers would add up to exactly \( B \).

Now add two more elements \( b_1 \) and \( b_2 \) to \( A \), where:

\[ s(b_1) = 2 \sum_{i=1}^{k} s(a_i) - B \]

\[ s(b_2) = \sum_{i=1}^{k} s(a_i) + B \]
Note that the sum of the sizes of all the elements in $A$ is:

\[ s(a_i) + s(b_1) + s(b_2) \]

\[ = \sum_{i=1}^{k} s(a_i) + (2\sum_{i=1}^{k} s(a_i) - B) + (\sum_{i=1}^{k} s(a_i) + B) \]

\[ = 4 \sum_{i=1}^{k} s(a_i) \]

Thus, any partition of $A$ into equal-sizes subsets $A'$ and $A - A'$, the sum of the sizes of the elements in those two partitions must both equal:

\[ 2 \sum_{i=1}^{k} s(a_i) \]

Furthermore, $b_1$ and $b_2$ cannot be in the same set.

\[ s(b_1) = 2 \sum_{i=1}^{k} s(a_i) - B \]

\[ s(b_2) = \sum_{i=1}^{k} s(a_i) + B \]

\[ = 3 \sum_{i=1}^{k} s(a_i) \]

1.79
Thus, in any partition \( A' \) of \( A \) into two equal-sized subsets, \( b_1 \) and \( b_2 \) must be in different sets:

\[
A' \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quarter
There are several “general” types of transformations:

- Restriction
- Local Replacement
- Component Design
Proving a problem $\Pi \in \text{NP}$ NP-complete by restriction consists of showing that $\Pi$ contains another known NP-complete problem $\Pi'$ as a special case.

What is a “special case” of a problem?
- Let $I$ be all instances of a problem $\Pi$, and let $I'$ be all instances of a problem $\Pi'$.
- If $I' \subseteq I$ and $I' \in Y_{\Pi'}$ IFF $I \in Y_{\Pi}$ then $\Pi'$ is said to be a special case of $\Pi'$.

Furthermore, if $\Pi$ and $\Pi'$ are both in NP, and if $\Pi'$ is NP-complete, then so is $\Pi$.

The heart of a proof by restriction is the specification of restrictions to be placed on the instances of $\Pi$ so that the resulting problem will be identical to $\Pi'$.

Note that a proof by restriction is still a transformation, i.e., it is not a different kind of NP-completeness proof.
DIRECTED HAMILTONIAN CIRCUIT (DHC)

INSTANCE: A directed graph $G = (V, E)$.

QUESTION: Does $G$ have a Hamiltonian circuit, that is an ordering $< v_1, v_2, ..., v_n >$ of the vertices of $G$, where $n = |V|$, such that $\{v_n, v_1\} \in E$ and $\{v_i, v_{i+1}\} \in E$ for all $i$, $1 \leq i < n$.

HAMILTONIAN CIRCUIT (HC)

INSTANCE: A Graph $G = (V, E)$.

QUESTION: Does $G$ have a Hamiltonian circuit, that is an ordering $< v_1, v_2, ..., v_n >$ of the vertices of $G$, where $n = |V|$, such that $\{v_n, v_1\} \in E$ and $\{v_i, v_{i+1}\} \in E$ for all $i$, $1 \leq i < n$.

Fact: HC is NP-complete.
**Theorem:** DHC is NP-complete

**Proof:**
1) DHC $\in$ NP
2) HC $\preceq$ DHC, by restriction.

*Proof:* Consider those instances of DHC where there is an edge from $u$ to $v$ if and only if there is an edge from $v$ to $u$.

- Recall that technically a restriction is, in fact, a transformation.

- In what sense is the above proof a transformation?
**Exact Cover by 3-Sets (X3C)**

**EXACT COVER BY 3-SETS (X3C)**

INSTANCE: A finite set $X$ with $|X| = 3q$ and a collection $C$ of 3-element subsets of $X$.

QUESTION: Does $C$ contain an exact cover for $X$, that is, a sub-collection $C' \subseteq C$ such that every element of $X$ occurs in exactly one member of $C'$?

How does this differ from 3DM?

**3-DIMENSIONAL MATCHING (3DM)**

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: Does $M$ contain a matching, that is, a subset $M' \subseteq M$ such that $|M'| = q$ and no two elements of $M'$ agree in any coordinate?
Theorem: X3C is NP-complete

Proof:
1) X3C ∈ NP
2) 3DM ∝ X3C

Let $M \subseteq W \times Y \times Z$, where $W$, $Y$ and $Z$ are disjoint sets having the same number $q$ of elements, be an instance of 3DM.

Construct an instance of X3C as follows:

Let $X = W \cup Y \cup Z$
Let $C = \{ \{a,b,c\} \mid (a,b,c) \in M \}$

It can easily be verified that $C$ contains an exact 3-cover for $X$ iff $M$ contains a matching.
**Theorem:** X3C is NP-complete

**Proof:**

1) X3C ∈ NP
2) 3DM ∝ X3C

Another way to view this is as a restriction from X3C to 3DM – consider those instances of X3C where X can be partitioned into 3 disjoint sets W, Y, and Z, and where each 3-element set C contains exactly one element from each of these sets.
MINIMUM COVER

INSTANCE: A collection \( C \) of subsets of a set \( S \), positive integer \( K \).

QUESTION: Does \( C \) contain a cover for \( S \) of size \( K \) or less, that is, a subset \( C' \subseteq C \) with \( |C'| \leq K \) and such that:

\[
\bigcup_{c \in C'} c = S \, ?
\]

Proof: Restrict to X3C by considering those instances having \( |c| = 3 \) for all \( c \in C \), and having \( K = |S|/3 \).

EXACT COVER BY 3-SETS (X3C)

INSTANCE: A finite set \( X \) with \( |X| = 3q \) and a collection \( C \) of 3-element subsets of \( X \).

QUESTION: Does \( C \) contain an exact cover for \( X \), that is, a sub-collection \( C' \subseteq C \) such that every element of \( X \) occurs in exactly one member of \( C' \)?
HITTING SET

INSTANCE: A collection $C$ of subsets of a set $S$, positive integer $K$.

QUESTION: Does $S$ contain a hitting set for $C$ of size $K$ or less, that is, a subset $S' \subseteq S$ with $|S'| \leq K$ and such that $S'$ contains at least one element from each subset in $C$?

Proof: Restrict to VC by considering those instances having $|c| = 2$ for all $c \in C$.

VERTEX COVER (VC)

INSTANCE: A Graph $G = (V, E)$ and a positive integer $K \leq |V|$.

QUESTION: Is there a vertex cover of size $K$ or less for $G$, that is, a subset $V' \subseteq V$ such that $|V'| \leq K$ and, for each edge $\{u, v\} \in E$, at least one of $u$ and $v$ belongs to $V'$?
SUBGRAPH ISOMORPHISM

INSTANCE: Two graphs $G = (V_1, E_1)$ and $H = (V_2, E_2)$.

QUESTION: Does $G$ contain a subgraph isomorphic to $H$, that is, a subset $V \subseteq V_1$ and a subset $E \subseteq E_1$ such that $|V| = |V_2|$, $|E| = |E_2|$, and there exists a one-to-one function $f : V_2 \rightarrow V$ satisfying $\{u, v\} \in E_2$ if and only if $\{f(u), f(v)\} \in E$.

Proof: Restrict to CLIQUE by considering those instances where $H$ is a complete graph, that is, $E_2$ contains all possible edges joining two members of $V_2$.

CLIQUE

INSTANCE: A Graph $G = (V, E)$ and a positive integer $J \leq |V|$.

QUESTION: Does $G$ contain a clique of size $J$ or more?
Let $G=(V,E)$ be a graph and let $E' \subseteq E$. Then $T=(V,E')$ is said to be a spanning tree if $T$ is connected and $|E'| = |V| - 1$ (i.e., $T$ contains no cycles).

**BOUNDED DEGREE SPANNING TREE**

**INSTANCE:** A graph $G = (V, E)$ and a positive integer $K \leq |V| - 1$.

**QUESTION:** Is there a spanning tree for $G$ in which no vertex has degree exceeding $K$, that is, a subset $E' \subseteq E$ such that $|E'| = |V| - 1$, the graph $G' = (V, E')$ is connected, and no vertex in $V$ is included in more than $K$ edges from $E'$?

**Proof:** Restrict to HAMILTONIAN PATH by considering only instances where $K=2$.

**HAMILTONIAN PATH (HP)**

**INSTANCE:** A Graph $G = (V, E)$, and vertices $u, v \in V$.

**QUESTION:** Is there a Hamiltonian path from vertex $u$ to vertex $v$, that is an ordering $< v_1, v_2, \ldots, v_n >$ of the vertices of $G$, where $n = |V|$, such that $u = v_1$, $v = v_n$, and $\{v_i, v_{i+1}\} \in E$ for all $i, 1 \leq i < n$.

**FACT:** HP is NP-complete (see page 60).
MINIMUM EQUIVALENT DIGRAPH

INSTANCE: A directed graph $G = (V, A)$ and a positive integer $K \leq |A|$.

QUESTION: Is there a directed graph $G' = (V, A')$ such that $A' \subseteq A$, $|A'| \leq K$, and such that, for every pair of vertices $u$ and $v$ in $V$, $G'$ contains a directed path from $u$ to $v$ if and only if $G$ contains a directed path from $u$ to $v$.

Proof: Restrict to DIRECTED HAMILTONIAN CIRCUIT by considering those instances where $G$ is strongly connected, that is, contains a path from every vertex $u$ to every vertex $v$, and $K = |V|$. Note that this is technically a restriction to DIRECTED HAMILTONIAN CIRCUIT FOR STRONGLY CONNECTED DIGRAPHS, but the NP-completeness of that problem follows immediately from the constructions given for HC and DIRECTED HC.
KNAPSACK

INSTANCE: A finite set $U$ and a “size” $s(u) \in \mathbb{Z}^+$ and a value $s(v) \in \mathbb{Z}^+$ for each $u \in U$, a size constraint $B \in \mathbb{Z}^+$, and a value goal $K \in \mathbb{Z}^+$.

QUESTION: Is there a subset $U' \subseteq U$ such that:

$$\sum_{u \in U'} s(u) \leq B \quad \text{and} \quad \sum_{u \in U'} v(u) \geq K$$

Proof: Restrict to PARTITION by considering only those instances where $s(u) = v(u)$ for all $u \in U$ and:

$$B = K = 1/2 \sum_{u \in U} s(u)$$

PARTITION

INSTANCE: A finite set $A$ and a “size” $s(a) \in \mathbb{Z}^+$ for each $a \in A$.

QUESTION: Is there a subset $A' \subseteq A$ such that:

$$\sum_{a \in A'} s(a) = \sum_{a \in A - A'} s(a)$$
Knapsack

**KNAPSACK**

INSTANCE: A finite set $U$ and a “size” $s(u) \in \mathbb{Z}^+$ and a value $s(v) \in \mathbb{Z}^+$ for each $u \in U$, and a value goal $K \in \mathbb{Z}^+$.

QUESTION: Is there a subset $U' \subseteq U$ such that:

$$\sum_{u \in U'} v(u) \geq K$$
MULTIPROCESSOR SCHEDULING

INSTANCE: A finite set $A$ of “tasks,” a “length” $l(a) \in \mathbb{Z}^+$ for each $a \in A$, a number $m \in \mathbb{Z}^+$ of “processors,” and a “deadline” $D \in \mathbb{Z}^+$.

QUESTION: Is there a partition $A = A_1 \cup A_2 \cup \ldots \cup A_m$ of $A$ into $m$ disjoint sets such that:

$$
\max \left\{ \sum_{a \in A_i} l(a) : 1 \leq i \leq m \right\} \leq D?
$$

Proof: Restrict to PARTITION by considering only those instances where $m=2$ and:

$$
D = 1/2 \sum_{a \in A} l(a)
$$

PARTITION

INSTANCE: A finite set $A$ and a “size” $s(a) \in \mathbb{Z}^+$ for each $a \in A$.

QUESTION: Is there a subset $A' \subseteq A$ such that:

$$
\sum_{a \in A'} s(a) = \sum_{a \in A-A'} s(a)
$$
Invalid Restriction

Subset Sum
INSTANCE: Integers $a_1, a_2, \ldots, a_n$ and integer $B$.
QUESTION: Is there a sequence of 0’s and 1’s, $x_1, x_2, \ldots, x_n$ such that:

$$\sum_{i=1}^{n} a_i x_i = B?$$

Fact: Subset Sum is NP-complete

Real Subset Sum
INSTANCE: Integers $a_1, a_2, \ldots, a_n$ and integer $B$.
QUESTION: Is there a sequence of real numbers $x_1, x_2, \ldots, x_n$ such that:

$$\sum_{i=1}^{n} a_i x_i = B?$$
Invalid Restriction, Cont.

- Claim (from a Ph.D dissertation):
  - Subset Sum is just a special case of Real Subset Sum
  - Therefore, Real Subset Sum is NP-complete

- The proof is in error since it restricts the question, and not the instances.

- In fact, the problem is actually trivially solvable in polynomial time simply taking $x_1 = B/a_1$ and $x_i = 0$, for all $i$, where $2 \leq i \leq n$.

- Also note that the “restriction” does not preserve yes and no instances (in fact, all are yes instances).
Some other invalid restrictions:
- 4-DM is NP-complete because it contains 3-DM as a special case.
- 4-SAT is NP-complete because it contains 3-SAT as a special case.

In these cases, the instances of one aren’t even a subset of the other!
Homework #2

- See page 75, the first group:
  - Longest Path
  - Set Packing
  - Partition into Hamiltonian Subgraphs
  - Largest Common Subgraph
  - Minimum Sum of Squares

- Give an NP-completeness proof for each of the above by restriction.

- Hand in any two.
Transformations are:
- sufficiently non-trivial to warrant spelling out details (in contrast to restriction)
- still relatively simple

First example – transforming SAT to 3-SAT:
- each clause in SAT was “locally replaced” by 1 or more clauses in 3-SAT
- Replacement of one clause was independent of all other clauses
Partition into Triangles

INSTANCE: A Graph $G = (V, E)$, with $|V| = 3q$.

QUESTION: Is there a partition of $V$ into $q$ disjoint sets $V_1, V_2, \ldots, V_q$ of three vertices each such that, for each $V_i = \{v_{i[1]}, v_{i[2]}, v_{i[3]}\}$ the three edges $\{v_{i[1]}, v_{i[2]}\}$, $\{v_{i[1]}, v_{i[3]}\}$ and $\{v_{i[2]}, v_{i[3]}\}$ all belong to $E'$?

Theorem: Partition into Triangles is NP-complete

Proof:

1) Partition into Triangles $\in$ NP

2) X3C $\propto$ Partition into Triangles
Let set \( X \) with \(|X| = 3q\) and let \( C \) be a collection of 3-element subsets of \( X \).

Construct a graph \( G = (V,E) \) as follows:

For each set \( c_i = \{x_i, y_i, z_i\} \) in \( C \), create the following vertices and edges:

That’s it!
Can you draw the graph?

\[ q = 3 \text{ (not part of input technically)} \]
\[ X = \{a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3\} \]

\[ M = \{(a_1, b_2, c_1), (a_2, b_1, c_3), (a_3, b_3, c_2), (a_1, b_1, c_3), (a_2, b_2, c_1)\} \]

Note the first 3 form an exact 3-cover.
The resulting graph $G = (V, E)$ can be partitioned into triangles IFF there is a sub-collection $C' \subseteq C$ such that every element of $X$ occurs in exactly one member of $C'$.

(if) Suppose there is a sub-collection $C' \subseteq C$ such that every element of $X$ occurs in exactly one member of $C'$. Consider each triple from $C$.

It is easy to see that this gives a partition of $G = (V, E)$ into triangles.
(only if) Suppose the resulting graph \( G = (V, E) \) can be partitioned into triangles.

- Note that each component \textbf{must} be partitioned into one of two ways.

  - In the first case, put the corresponding triple in \( C' \), and in the second case, don’t.

  - It is easy to see that the resulting sub-collection \( C' \subseteq C \) is an exact cover for \( X \).
The most complicated of the 3 types of proofs.

Constructs “components” in the target instance from “components” of the given instance:
- not a simple one-to-one mapping, in contrast to local replacement
- typically between different types of problems

Components serve two purposes typically:
- making choices – selecting vertices or truth assignments
- testing properties – making sure all edges are covered or that clauses are satisfied

Additionally, the choice-components are connected to the testing components in such a way to ensure choices satisfy required properties.

Examples:
- 3 Dimensional Matching (slide 27)
- Vertex Cover (slide 40) – or is this local replacement?
- Hamiltonian Circuit (slide 49)
MINIMUM TARDINESS SEQUENCING

INSTANCE: A finite set $T$ of “tasks,” each $t \in T$ having “length” 1 and a “deadline” $d(t) \in \mathbb{Z}^+$, a partial order $<$ on $T$, and a non-negative integer $K \leq |T|$.

QUESTION: Is there a “schedule” $\sigma: T \rightarrow \{0, 1, \ldots, |T|-1\}$ such that $\sigma(t) \neq \sigma(t')$ whenever $t \neq t'$, such $\sigma(t) < \sigma(t')$ whenever $t < t'$, and such that $|\{t \in T: \sigma(t)+1 > d(t)\}| \leq K$?

**Theorem:** Minimum Tardiness Sequencing is NP-complete

**Proof:**

1) Minimum Tardiness Sequencing $\in$ NP

2) CLIQUE $\propto$ Minimum Tardiness Sequencing

Let the graph $G=(V,E)$ and positive integer $J \leq |V|$ be an arbitrary instance of CLIQUE.

Construct an instance of Minimum Tardiness Sequencing as follows.
Component Design

- Tasks:
  \[ T = V \cup E \]

- Task Partial Order:
  \[ t < t' \iff t \in V, t' \in E, \text{ and } t \text{ is an endpoint of edge } t' \]

- Task Deadlines:
  \[ d(t) = \begin{cases} 
  \frac{l(j+1)}{2} & \text{if } t \in E \\
  |V| + |E| & \text{if } t \in V 
  \end{cases} \]

- Tardy-Task Limit:
  \[ K = |E| - (J(J-1)/2) \]
Observation #1: A schedule is an ordering of the tasks, from 0 to $|T|-1$.

Consequently, any schedule for the tasks will look as follows:

We need to show that the $G = (V,E)$ contains a clique of size $J$ IFF the resulting set of tasks has a minimum tardiness schedule.
(only if) Suppose the graph $G$ contains a clique of size $J$.

Construct a schedule as follows:

- `clique vertices`: $J$ vertex tasks
- `clique edges`: $J(J-1)/2$ edge tasks
- `|V| - J`: vertex tasks
- `|E| - J(J-1)/2`: edge tasks

$T = |V| + |E|$
(if) Suppose there is a “schedule” \( \sigma : T \rightarrow \{0, 1, \ldots, |T| - 1\} \) such that \( \sigma(t) \neq \sigma(t') \) whenever \( t \neq t' \), such \( \sigma(t) < \sigma(t') \) whenever \( t < t' \), and such that \(|\{t \in T : \sigma(t) + 1 > d(t)\}| \leq K?\)

Observation #1: At least \( J(J-1)/2 \) of the edge tasks must be scheduled before the edge deadline; otherwise, too many edge tasks would be late.

<table>
<thead>
<tr>
<th>( \geq J ) vertex tasks</th>
<th>( \geq J(J-1)/2 ) edge tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( J(J+1)/2 )</td>
</tr>
</tbody>
</table>

\( T = |V| + |E| \)

Observation #2: At least \( J \) vertex tasks must also occur prior to the edge task deadline; this is because the minimum number of vertices on \( J(J-1)/2 \) edges is \( J \).
Observation #3: Since there are at least $J$ vertex tasks, and at least $J(J-1)/2$ edge tasks prior to the deadline $J(J+1)/2$, it follows there must be exactly $J$ vertex tasks and exactly $J(J-1)/2$ edge tasks prior to the deadline $J(J+1)/2$.

Observation #4: The $J$ vertex tasks must also occur before the $J(J-1)/2$ edge tasks because of the partial order $\leq$.

It follows that the $J$ vertex tasks and the $J(J-1)/2$ edge tasks form a clique in the graph $G=(V,E)$. 
Which category would these be?

- 3-SAT (slide 10)
- Hypergraph 2-colorability (slide 22)
- Clique (slide 45)
- Independent Set (slide 45)
- Independent Set (slide 68)
- 4-Dimensional Matching (71)
- Partition (slide 75)