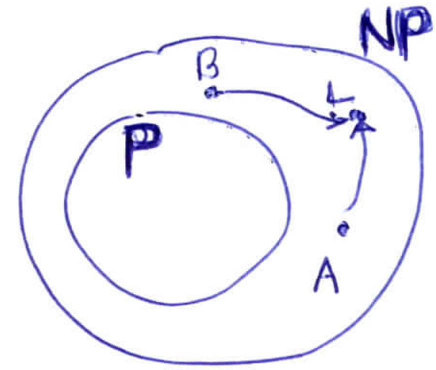
# NP-Completeness




**Def** A language L is called NP-complete if

① $L \in NP$

② $\forall$ language $A \in NP$, $A \leq_p L$.

**Theorem** The following language is NP-complete

$L = \{ \langle M, x, \#^k \rangle \mid$ M is a NTM that has accepting computation on $x$ with $\leq k$ steps $\}$

**Proof**

① $L \in NP$ since it is accepted by

$M_{sim}$ := " On input $\langle M, x, \#^k \rangle$,
Simulate M on $x$ for at most k steps.
If M accepts, accept
else reject "

Note: Since M is NTM, so is $M_{sim}$.

$\langle M, x, \#^k \rangle \in L \Rightarrow M$ has accepting comput$^n$ on $x$ w/ $\leq k$ steps

$\Rightarrow M_{sim}$ has an accepting comput$^n$.

$\langle M, x, \#^k \rangle \notin L \Rightarrow M$ has no accepting comput$^n$ on $x$ w/ $\leq k$ steps.

$\Rightarrow M_{sim}$ has no accepting comput$^n$.

Further $M_{sim}$ runs in time $\text{poly}(|\langle M \rangle|, |x|, k)$.

② Now we show that $\forall A \in NP,\ A \leq_p L$.

Let $A$ be accepted by polytime NTM $M_A$.

Here is the polytime reduction from $A$ to $L$!

$$A \longrightarrow L$$

$$x \rightsquigarrow \langle M_A, x, \#^{n^c} \rangle$$

where $|x| = n$ <u>and</u> $M_A$ has runtime $n^c$.

$x \in A \iff M_A$ has an accepting comput$^n$ on $x$ w/ $\leq n^c$ steps, $|x| = n$

$\iff \langle M_A, x, \#^{n^c} \rangle \in L$.

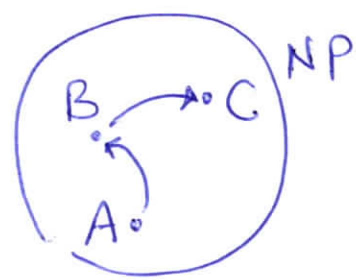
This shows that NP-complete languages exist. It turns out that many "natural" languages/problems are NP-complete!

<u>Theorem</u> [Cook-Levin] 3SAT is NP-complete.

<u>Theorem</u> To show that a language $C$ is NP-complete, one shows that

① $C \in NP$.

② For <u>some</u> NP-complete language $B$, $B \leq_p C$.



<u>Proof</u> ① Given, $C \in NP$.

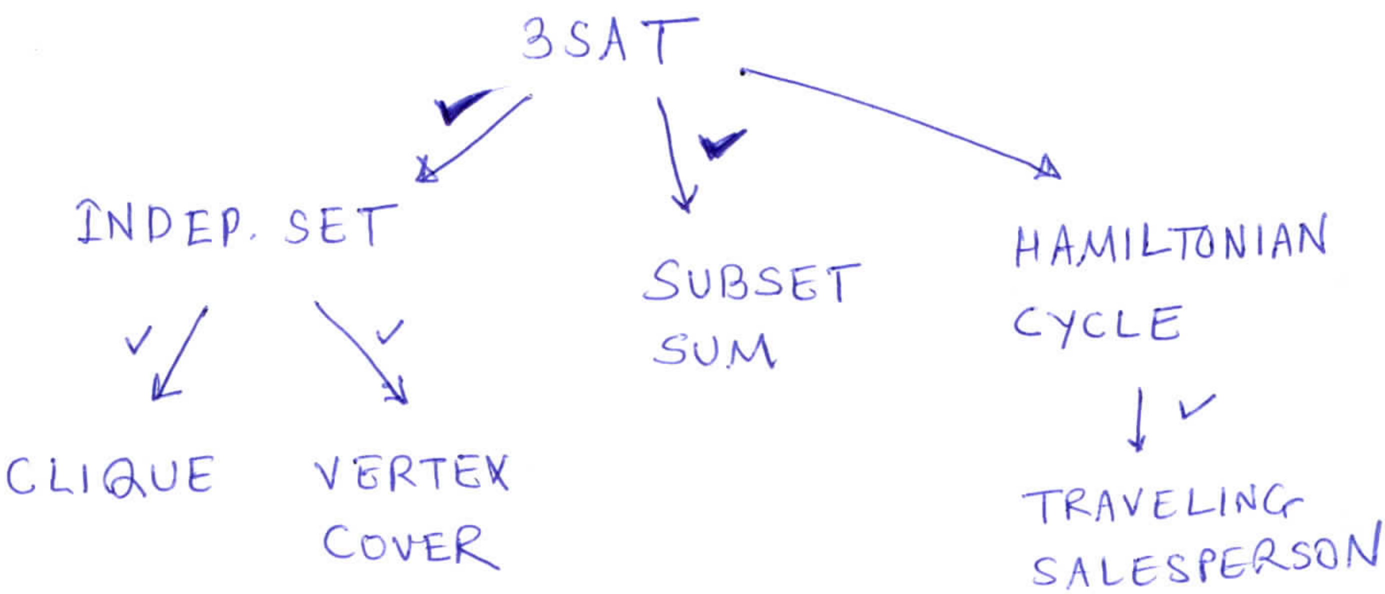
② We need to show that $\forall A \in NP$, $A \le_p C$.

Since $B$ is NP-complete, $A \le_p B$.

Given, $B \le_p C$.

Hence $A \le_p C$. ▨

Now we show that the following problems are NP-complete.



# 3SAT

Example of a 3SAT instance:

$$(x \vee y \vee z) \wedge (x \vee \bar{y} \vee \bar{w}) \wedge (\bar{y} \vee u \vee v) \wedge$$
$$(\bar{z} \vee \bar{u} \vee \bar{v}) \wedge (\bar{x} \vee z \vee w).$$

Def A 3CNF/3SAT formula $\phi$ is

$$\phi = C_1 \wedge C_2 \wedge C_3 \cdots \wedge C_m$$

where each clause $C_r$, $1 \le r \le m$, is of form $C_r = l_i \vee l_j \vee l_k$, $1 \le i,j,k \le n$

Here $x_1, x_2, \ldots, x_n$ are Boolean variables and $l_i = x_i$ or $l_i = \bar{x}_i$ are literals.

Def A 3CNF formula $\phi$ has a satisfying assignment ($\phi$ is satisfiable) if there is an assignment

$$\sigma : \{x_1, x_2, \ldots x_n\} \to \{\overset{1}{\text{True}}, \overset{0}{\text{False}}\}$$

that makes $\phi$ evaluate to True (i.e. makes every clause $C_r$ True).

3SAT = $\{\langle\phi\rangle \mid \phi$ has a satisfying assignment$\}$.

Note 3SAT $\in$ NP. Following polytime NTM accepts it.

M := " On input $\langle\phi\rangle$,

Let $x_1, \ldots, x_n$ be variables of $\phi$.

Guess an assign. $\sigma : \{x_1, .., x_n\} \to \{T, F\}$

Accept if $\sigma$ satisfies $\phi$.

Reject otherwise. "

Theorem 3SAT is NP-complete.

Note 3SAT can be thought of as the language as above

or equivalently as the problem of deciding, given instance $\phi$, whether $\phi$ has a satisfying assign.

# INDEPENDENT SET

**Def** In a graph $G(V,E)$, an independent set $I \subseteq V$ is a subset s.t.
$\forall\, a, b \in I$, $(a,b) \notin E$.

**Def**
INDEP. SET $= \{ \langle G, k \rangle \mid$ G is a n-vertex graph that has an independent set of size $\geq k$ $\}$.

**Theorem** INDEP. SET is NP-complete.

**Proof** ① INDEP. SET $\in$ NP. Following polytime NTM accepts it.

M := " Given $\langle G, k \rangle$, $G = G(V,E)$,
Non-det. select a subset $I \subseteq V$.
Accept if $|I| \geq k$ and $I$ is an independent set.
Reject otherwise. "

② We show that 3SAT reduces to INDEP. SET.

$$3SAT \longrightarrow \text{INDEP. SET}$$

$$\phi \rightsquigarrow \langle G, k \rangle$$

S.t. $\phi \in 3SAT \iff \langle G, k \rangle \in \text{INDEP. SET}$.

We need to show that

ⓐ $\phi \in 3SAT \Rightarrow \langle G, k \rangle \in \text{INDEP. SET}$.

ⓑ $\langle G, k \rangle \in \text{INDEP. SET} \Rightarrow \phi \in 3SAT$.

Reduction

$\phi$.

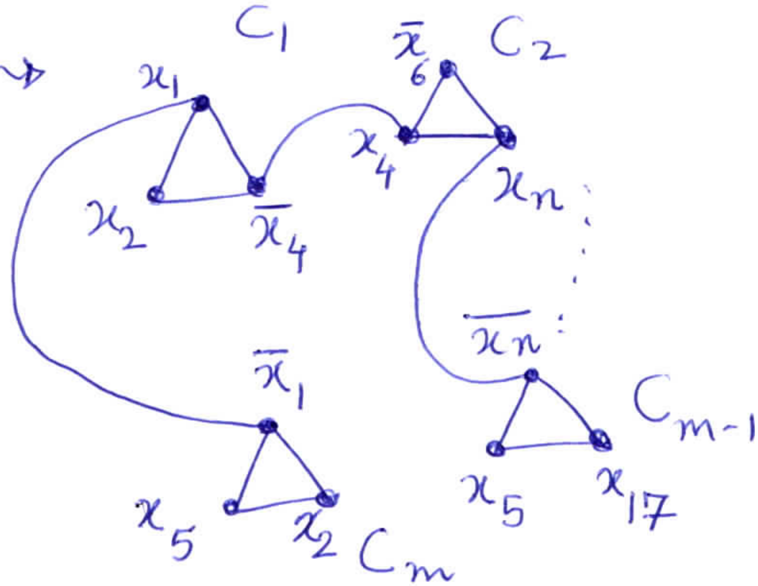
vars: $x_1, x_2, \ldots, x_n$

clauses: $C_1, C_2, \ldots, C_m$

$\rightsquigarrow$

G

$\boxed{k = m}$




- For each clause, construct a triangle whose vertices are labeled by the three literals in that clause.

- In addition, every pair of vertices labeled as $x_i$ and $\bar{x}_i$ are connected by an edge.
- This yields the graph $G$. Let $k = m$.

(a) $\phi$ has a satisfying assignment $\Rightarrow$ $G$ has an indep set of size $m$.

<u>Proof</u> Let $\sigma : \{x_1, .., x_n\} \to \{T, F\}$ be a satisfying assignment. For every $i$, $1 \le i \le n$, either $x_i$ or $\bar{x}_i$ is set to True (but not both).

Since $\sigma$ satisfies $\phi$, every clause $C_r$ contains a True literal.

Let $I$ = Subset of vertices obtained by picking one vertex from each clause/triangle $C_r$ that is labeled by a True literal.

clearly $|I| = m$.

Moreover, $I$ is indep. set since it consists of only True literals and hence does not contain both $x_i$ and $\bar{x}_i$ for any $1 \le i \le n$.

(b) G has an indep set of size $m$ $\Rightarrow$ $\phi$ has a satisfying assignment.

Proof Note first that the said independent set $I$ of size $m$ must contain exactly one vertex from each triangle/clause.

Declare all literals (labels of vertices in $I$) in $I$ to be True. This defines an assignment $\sigma$ to the variables $\{x_1, \dots, x_n\}$ in an unambiguous manner since $I$ does not contain a pair $(x_i, \bar{x}_i)$.

Since $I$ contains one vertex from each triangle, each clause contains a True literal.

Hence $\sigma$ is a satisfying assignment to $\phi$.

▨

# CLIQUE

**Def** A clique in a graph $G(V,E)$ is a subset $C \subseteq V$ s.t.

$\forall\, a,b \in C,\ a \neq b,\ (a,b) \in E.$

**Def** $\bar{G}$ is the complement graph where

- $\bar{G}$ has the same vertex set $V$ as $G(V,E)$
- $\forall\, a,b \in V,\ a \neq b,$

  $(a,b)$ is an edge in $\bar{G}$ $\Leftrightarrow$ $(a,b)$ is not an edge in $G$.

**Theorem** CLIQUE is NP-complete where

CLIQUE $= \{ \langle G, k \rangle \mid$ G has a clique of size $\geq k$. $\}$.

**Proof** ① CLIQUE $\in$ NP.

Blah Blah Blah ....

② INDEPENDENT-SET reduces to CLIQUE.

**Theorem** INDEP. SET $\leq_p$ CLIQUE

**Proof** INDEP. SET $\longrightarrow$ CLIQUE

$\langle G', k' \rangle \rightsquigarrow \langle G, k \rangle$.

Let $G = \overline{G'}$, $k = k'$.

Self-evident that

$G'$ has indep. set of size $k'$ $\iff$ $G = \overline{G'}$ has a clique of size $k' = k$.

$\blacksquare$

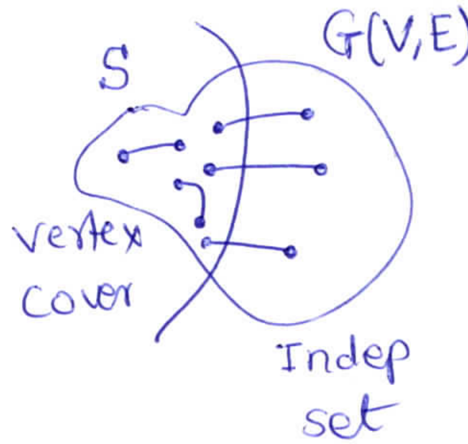
## VERTEX. COVER

**Def** A vertex cover in a graph $G(V, E)$ is a subset $S \subseteq V$ s.t.

$\forall (a,b) \in E$, either $a \in S$ or $b \in S$.

**Fact** $S$ is a vertex cover

$\iff V \setminus S$ is an indep. set.



**Theorem** VERTEX-COVER is NP-complete.

$$\text{VERTEX-COVER} = \{ \langle G, k \rangle \mid G \text{ has a vertex cover of size } \le k. \}$$

**Proof** 1 VERETEX-COVER $\in$ NP.

Blah Blah Blah ----

2 INDEPENDENT-SET $\le_p$ VERTEX-COVER.

$$\langle G', k' \rangle \rightsquigarrow \langle G, k \rangle.$$

$G'$ has independent set of size $k'$ $\iff$ $G$ has a vertex cover of size $k$.

**Reduction** Let $G = G'$.

$k = n - k'$ $\quad n = |G'|$.

Self-evident that

$G'$ has independent set $I$ of size $k'$ $\iff$ $G = G'$ has vertex cover $V \setminus I$ of size $n - k' = k$.

$\blacksquare$

**Theorem** HAMILTONIAN_CYCLE is NP-complete.

HAM.CYCLE = $\{ \langle G \rangle \mid G$ has a Hamiltonian cycle$\}$.

We skip the proof (= reduction from 3SAT).

**Theorem** T.S.P. is NP-complete.

T.S.P. = $\{ \langle G, wt, \ell \rangle \mid$ G is a complete graph with wt on the edges & has a tour of length $\leq \ell \}$

**Proof** (1) T.S.P. $\in$ NP.

Blah Blah Blah ...

(2) HAMILTONIAN-CYCLE $\leq_p$ T.S.P.

$G' \rightsquigarrow \langle G, wt, \ell \rangle$.

$G'(V', E')$

**Reduction** - Let G be complete graph on same vertex set $V'$.

- $wt(e) = \begin{cases} 1 & \text{if } e \in E' \\ 2 & \text{if } e \notin E' \end{cases}$

- $\ell = n$.

Now we show that

$G'$ has Hamiltonian cycle $\iff$ $\langle G, \underline{wt} \rangle$ has a tour of length $\leq n$.

ⓐ $\Rightarrow$ : The Hamiltonian cycle in $G'$ serves as a tour in $\langle G, \underline{wt} \rangle$; its length is $n$ since all its edges are edges of $G'$.

ⓑ $\Leftarrow$ : Consider a tour in $\langle G, \underline{wt} \rangle$ of length $\leq n$. Since it has $n$ edges and all edge weights are 1 or 2, the edges of the tour all must have weight 1. Hence the tour must correspond to a Hamiltonian cycle in $G'$.

∎

# SUBSET- SUM

$$\text{SUBSET-SUM} = \left\{ (a_1, \ldots, a_n ; t) \;\middle|\; \begin{array}{l} a_1, \ldots, a_n, t \text{ are non-neg} \\ \text{integers and} \\ \exists S \subseteq \{1, \ldots, n\},\ \sum_{i \in S} a_i = t \end{array} \right\}.$$

Note $a_1, \ldots, a_n, t$ are represented in binary or decimal

Theorem SUBSET- SUM is NP complete.

Proof ① SUBSET- SUM $\in$ NP

Blah Blah Blah ....

② We show that 3SAT $\leq_p$ SUBSET- SUM.

- Let $\phi$ be a 3SAT formula.

  vars: $x_1, \ldots, x_n$

  clauses: $C_1, C_2, \ldots, C_m$

- The SUBSET- SUM instance we construct
  - has integers in decimal.
  - no carries while adding integers.
  - digits not ~~shown~~ described are 0.

$$
\begin{array}{c|cccc|cccc|}
 & & & & & C_1 & C_2 & \cdots & C_m \\
\hline
x_1 & 1 & & & & 1 & 1 & & \\
\bar{x}_1 & 1 & & & & & & & \\
x_2 & & 1 & & & & 1 & & \\
\bar{x}_2 & & 1 & & & 1 & & & \\
\vdots & & & & & & & & \\
x_n & & & \cdots & 1 & & 1 & & \\
\bar{x}_n & & & & 1 & 1 & & & \\
\hline
g_1 & & & & & 1 & & & \\
h_1 & & & & & 1 & & & \\
g_2 & & & & & & 1 & & \\
h_2 & & & & & & 1 & & \\
\vdots & & & & & & & \ddots & \\
g_m & & & & & & & & 1 \\
h_m & & & & & & & & 1 \\
\hline
\end{array}
$$

$$
t \quad \begin{array}{|cccc|cccc|}
\hline
1 & 1 & \cdots & 1 & 3 & 3 & \cdots & 3 \\
\hline
\end{array}
$$

- rows represent the (decimal) integers.
- For illustration : $C_1 : x_1 \vee \bar{x}_2 \vee \bar{x}_n$

  $C_2 : x_1 \vee x_2 \vee x_n$

The SUBSET-SUM instance is as follows:

- Top-right block: an entry is 1 if the literal belongs to the clause and 0 otherwise.

- Remaining three blocks: as shown. and the target $t$

- Note: Only way to pick a subset of rows that sum to $t$ is to
  - pick exactly one of the two rows labeled $x_i$ or $\bar{x}_i$, for every $1 \leq i \leq n$.
  - pick none, one, or both of the rows labeled $g_j$, or $h_j$, for every $1 \leq j \leq m$.

Now we prove that

$\phi$ has a satisfying assignment $\Leftrightarrow$ There is a subset of rows that sums to $t$.

proof of $\Rightarrow$ :

- Let $\sigma$ be a satisfying assignment.
- For $1 \leq i \leq n$, between rows labeled $x_i$ or $\overline{x_i}$, pick the one corresponding to the True literal.
- For each clause $C_j$, at least one ~~True~~ of its literals is True and hence one, two, or three rows corresponding to its literals have been picked.
- Thus (in top-right block), column for $C_j$ has sum 1, 2, or 3.
- Depending on these three cases, we take both, only $g_j$, or none from the rows $g_j, h_j$ so that the sum in that column is exactly 3.

## Proof of $\Leftarrow$

Suppose there is a subset $S$ of rows that sums to $t$.

As noted, for every $i$, $1 \leq i \leq n$, exactly one of $x_i$ or $\bar{x}_i$ row is in $S$, and declare that literal to be True. This defines an assignment $\sigma$ to $x_1, x_2, \ldots, x_n$.

To show that $\sigma$ satisfies each clause $C_j$, we note that $C_j$ must contain a True literal. Otherwise (in the top right block), the sum in column $C_j$ is zero and even if one were to take both rows $g_j, h_j$, one wouldn't reach the sum 3 in that column. $\blacksquare$