CSE 6140/ CX 4140:

Computational Science and Engineering ALGORITHMS

Instructor: Anne Benoit

Visiting Associate Professor, CSE

Based on slides by Bistra Dilkina

Some NP-Complete Problems

- Six basic genres of NP-complete problems and paradigmatic examples.
  - Packing problems: SET-PACKING, INDEPENDENT SET.
  - Covering problems: SET-COVER, VERTEX-COVER.
  - Constraint satisfaction problems: SAT, 3-SAT.
  - Sequencing problems: HAMILTONIAN-CYCLE, TSP.
  - Partitioning problems: 3-COLOR, 3D-MATCHING.
  - Numerical problems: 2-PARTITION, SUBSET-SUM, KNAPSACK.

- In practice: Most NP problems are either known to be in P or NP-complete.

Independent set

Independent set (IS)

- Given a graph G=(V,E), find the largest independent set: a set of vertices in the graph with no edges between them.
- Decision version: is there an independent set of at least K vertices?

IS is in NP (Step 1)

- Certificate: set of vertices S
- Certifier: Check size of S ≥ K, and no pair of vertices in S is connected by an edge, O(n+m)

- Reduction by gadget (Step 2: choosing 3SAT)

3SAT → independent set of size K=m clauses in appropriately constructed graph G

3-Satisfiability Reduces to Independent Set

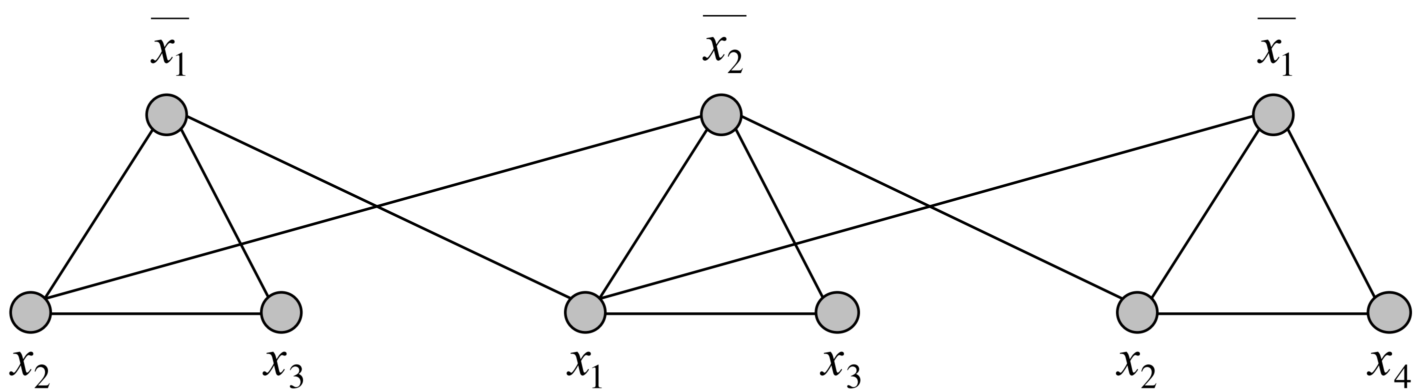
Claim. 3-SAT $\le_P$ INDEPENDENT-SET.

Pf. Given an instance Φ of 3-SAT (I_1), we construct an instance (G, k) of INDEPENDENT-SET (I_2) that has an independent set of size k iff Φ is satisfiable.

Construction (Step 3a)

- G contains 3 vertices for each clause, one for each literal.
- Connect 3 literals in a clause in a triangle.
- Connect literal to each of its negations.
- *The size of I_2 is polynomial in the size of I_1*

G



k = 3

$$\Phi = \left(\overline{x_1} \vee x_2 \vee x_3\right) \wedge \left(x_1 \vee \overline{x_2} \vee x_3\right) \wedge \left(\overline{x_1} \vee x_2 \vee x_4\right)$$

3 Satisfiability Reduces to Independent Set

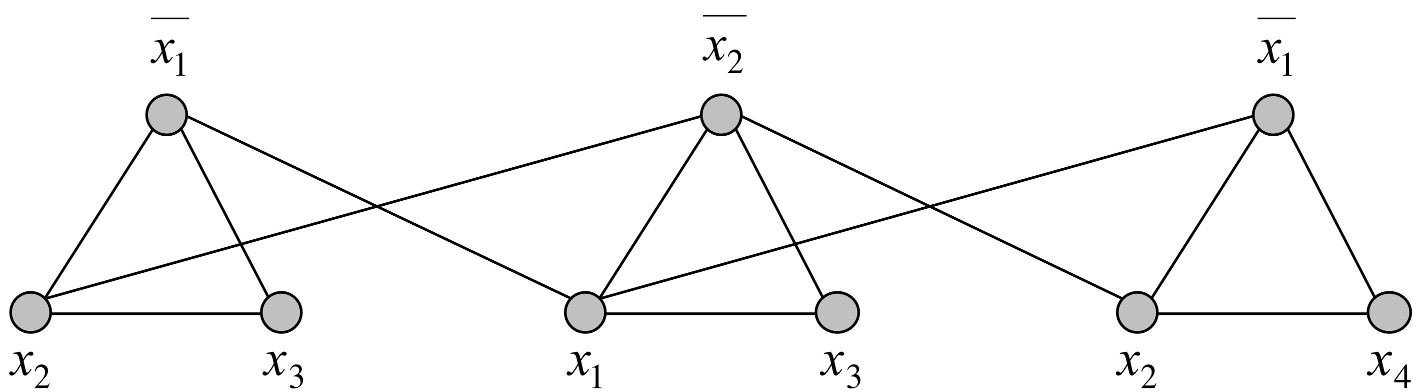
Claim. Φ is satisfiable iff G contains independent set of size $k = |\Phi|$
(I_1 has a solution $\Leftrightarrow$ I_2 has a solution)

(3b) => Given satisfying assignment (sol to I_1), select one true literal from each triangle. This is an indep. set of size k, hence a sol to I_2.

(3c) <= Let S be independent set of size k (sol. to I_2)

- S must contain exactly one vertex in each triangle.
- Set these literals to true. ← and any other variables in a consistent way
- Truth assignment is consistent and all clauses are satisfied.

G

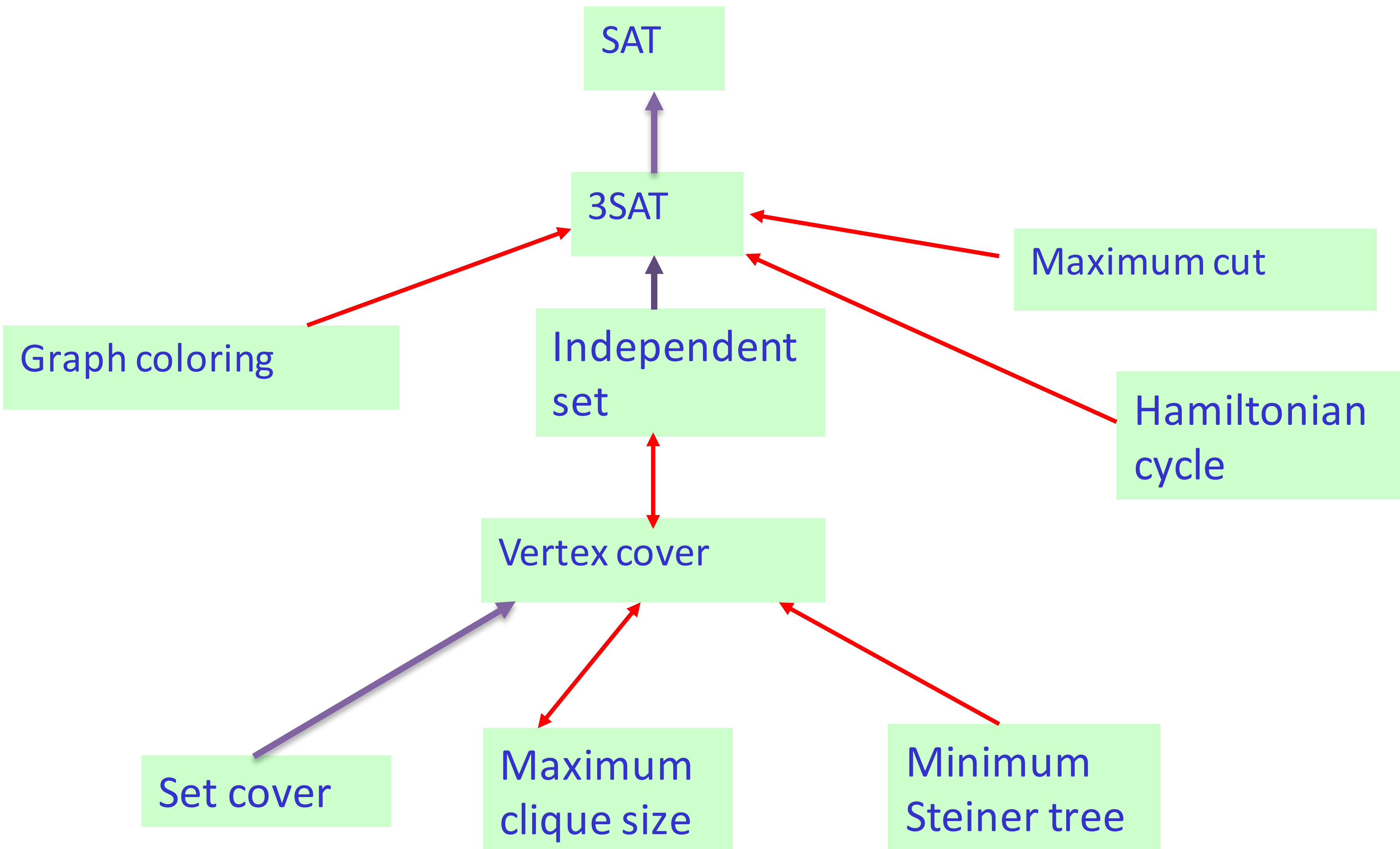


k = 3

$$\Phi = \left(\overline{x_1} \lor x_2 \lor x_3\right) \land \left(x_1 \lor \overline{x_2} \lor x_3\right) \land \left(\overline{x_1} \lor x_2 \lor x_4\right)$$

Examples of NP-complete problems

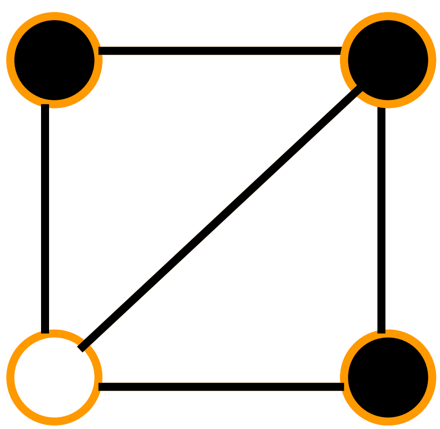
Summary of some NPc problems



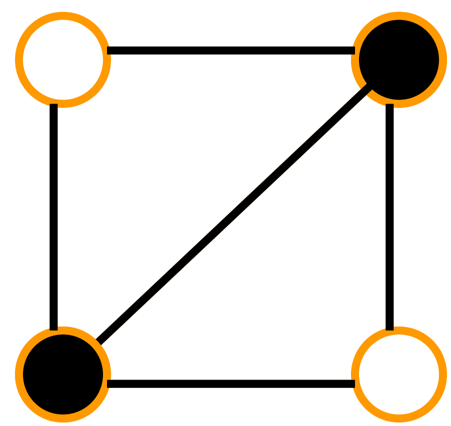
Vertex Cover

Vertex cover (VC)

- Given a graph G=(V,E), find the *smallest* number of vertices that cover *each edge*
- Decision problem: is there a set of at most K vertices that cover *each edge*?



vertex cover
of size 3



vertex cover
of size 2

Vertex Cover Decision Problem

- VC(G,k): Given a graph G and an integer K, does G have a vertex cover of size at most K?

- **Theorem: VC is NP-complete.**
- Proof:
- 1) Show that VC is in NP:
  - Certificate: a subset V' of the vertices
  - Certifier: check in polynomial time O(n+m) if |V'| ≤ K and if every edge has at least one endpoint in V'.

vertex cover in G of size k $\Leftrightarrow$ independent set in G of size |V| - k

Vertex Cover and Independent Set (KT 8.1)

Claim. We show a set of vertices S is an independent set of G iff $V - S$ is a vertex cover of G.

$\Rightarrow$

- Let S be any independent set.
- Consider an arbitrary edge (u, v).
- S independent $\Rightarrow u \notin S$ or $v \notin S$ $\Rightarrow$ $u \in V - S$ or $v \in V - S$.
- Thus, $V - S$ covers (u, v).

$\Leftarrow$

- Let $V - S$ be any vertex cover.
- Consider two nodes $u \in S$ and $v \in S$.
- Observe that $(u, v) \notin E$ since $V - S$ is a vertex cover.
- Thus, no two nodes in S are joined by an edge $\Rightarrow$ S independent set.

Claim. Solving IS(G,k) is equivalent to solving VC(G,n-k) and hence $VC \leq_P IS$ and $IS \leq_P VC$.

Vertex Cover is NP-complete

Independent Set: IS(G,k)

- Given a graph G=(V,E), find the largest independent set: a set of vertices in the graph with no edges between them.
- Decision version: is there an independent set of at least k vertices?

Vertex Cover: VC(G,k)

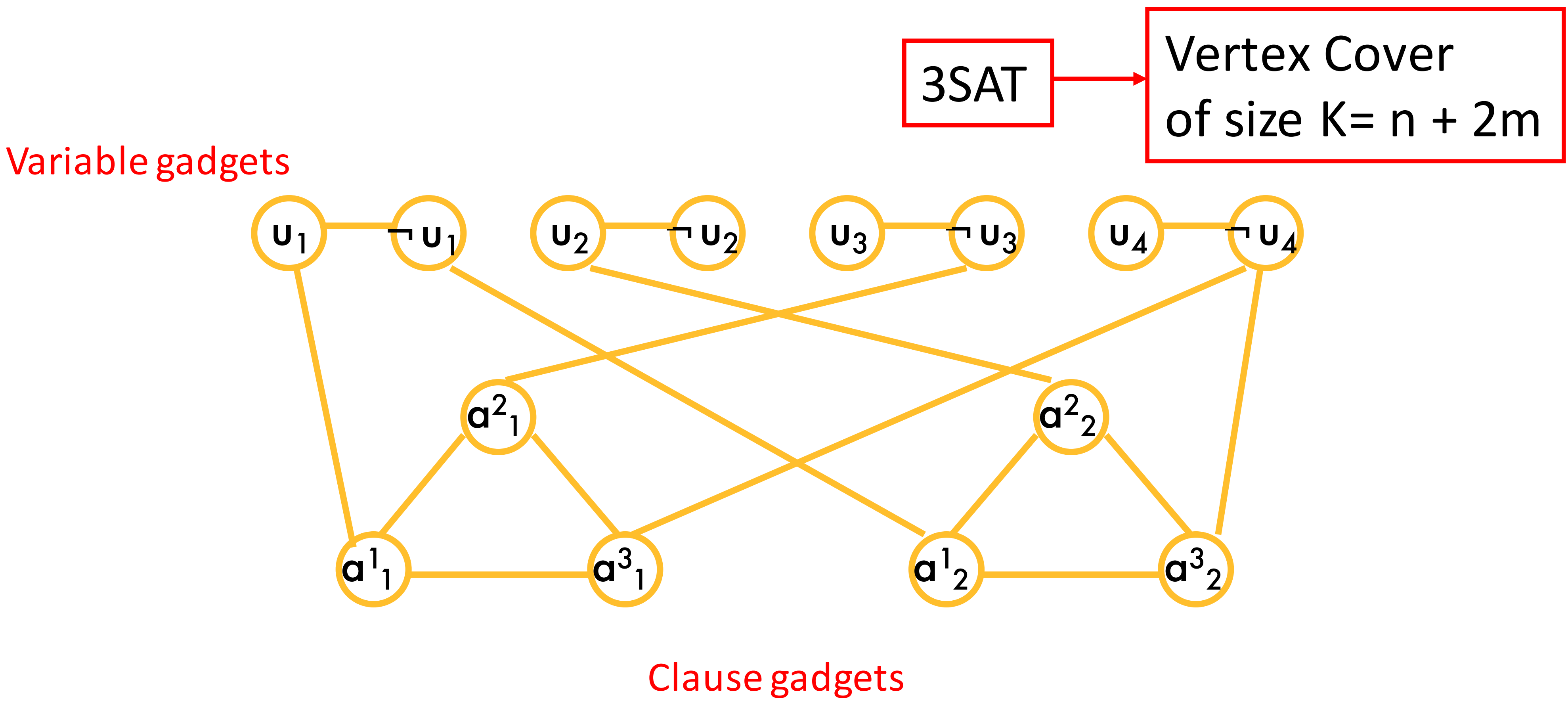
- Given a graph G and an integer k, does G have a vertex cover of size at most k?

VC is NP-complete because we showed:

- VC is in NP
- IS is NP-complete and IS $\leq_P$ VC, hence VC is NP-hard
- (Given IS(G,k), reduce it solving VC(G'=G,k'=|V|-k), correctness by proof on previous slide)

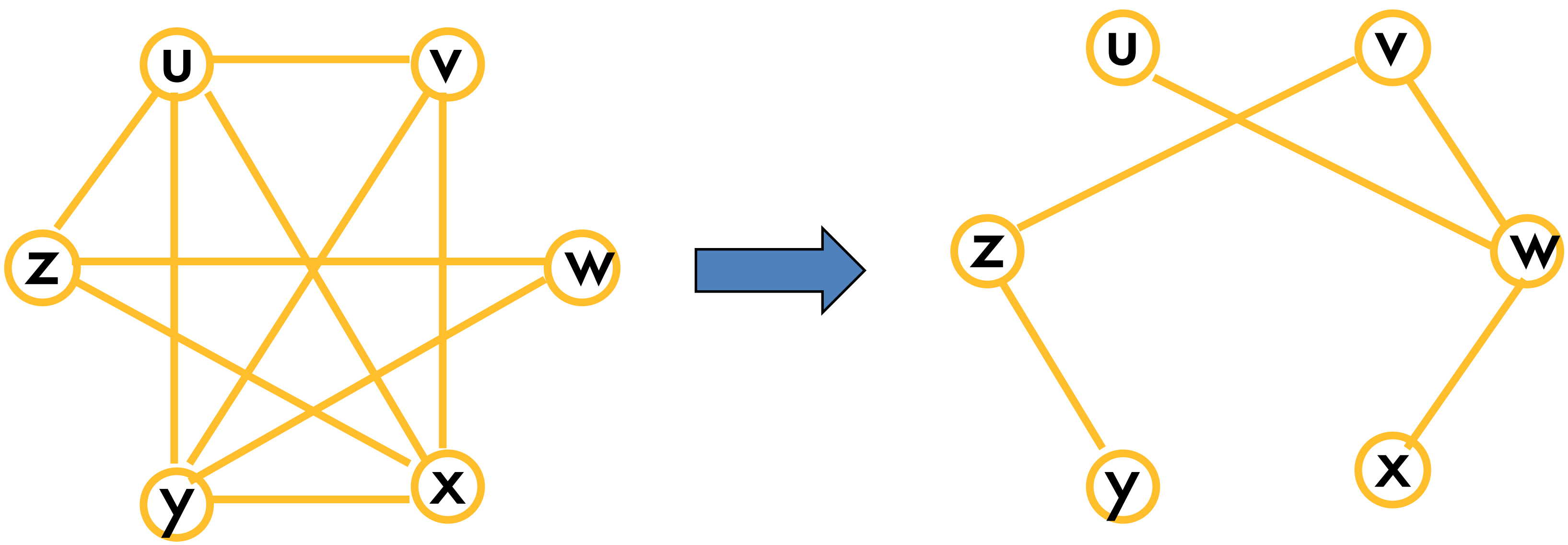
Vertex Cover

- Vertex cover (VC)
  - Given a graph G=(V,E) and an integer K, is there a set of at most K vertices that cover *each edge*?
    - By gadget (similar to 3SAT to IS)



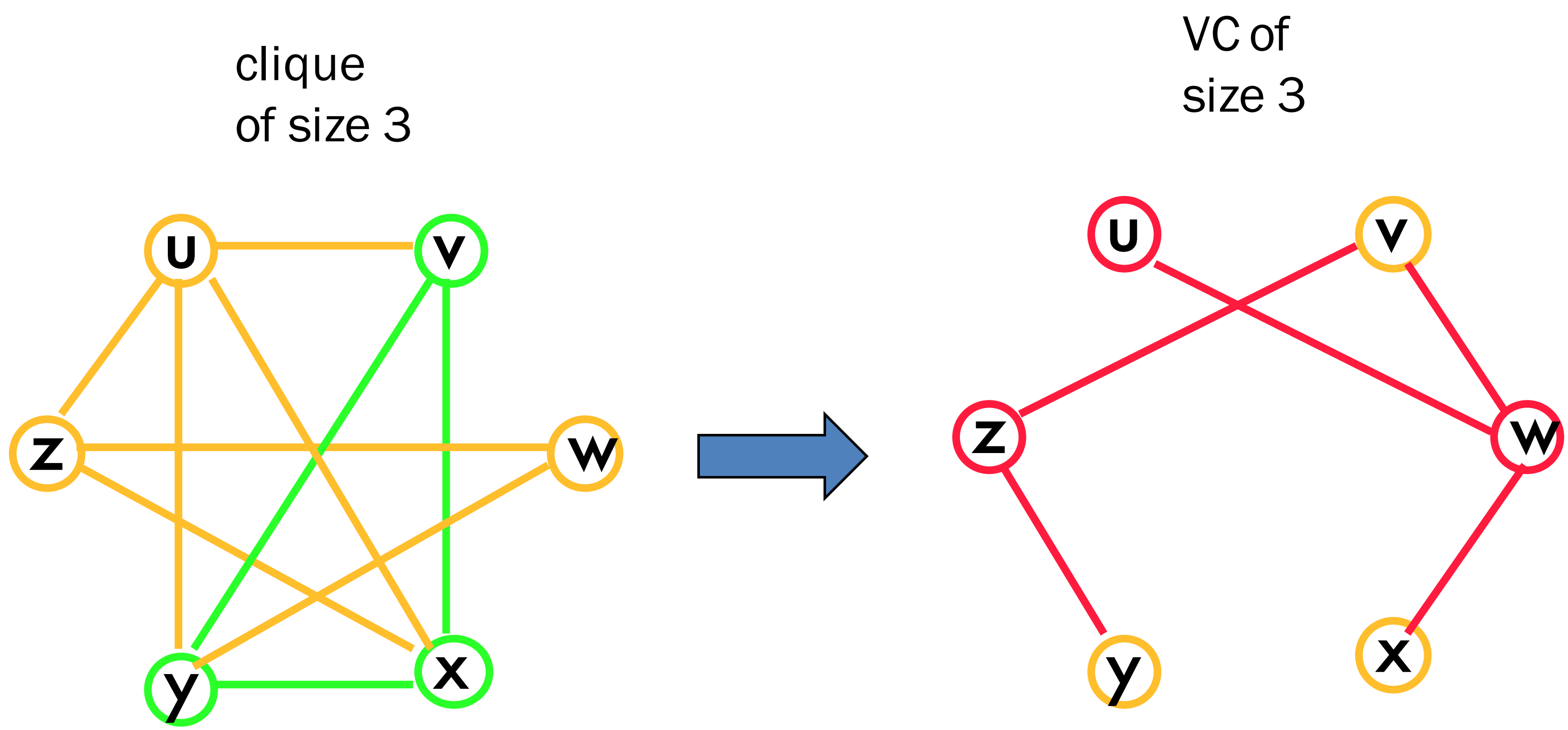
CLIQUE vs. VC (simple equivalence)

- CLIQUE(G,k): does G contain a completely connected subgraph of size at least K?
- The complement of graph G = (V,E) is the graph G_c = (V,E_c), where E_c consists of all the edges that are missing in G.
- CLIQUE(G,k) equivalent to VC(G_c,n-k)



CLIQUE vs. VC

- **Theorem**: V' is a clique of G if and only if V – V' is a vertex cover of G_c.



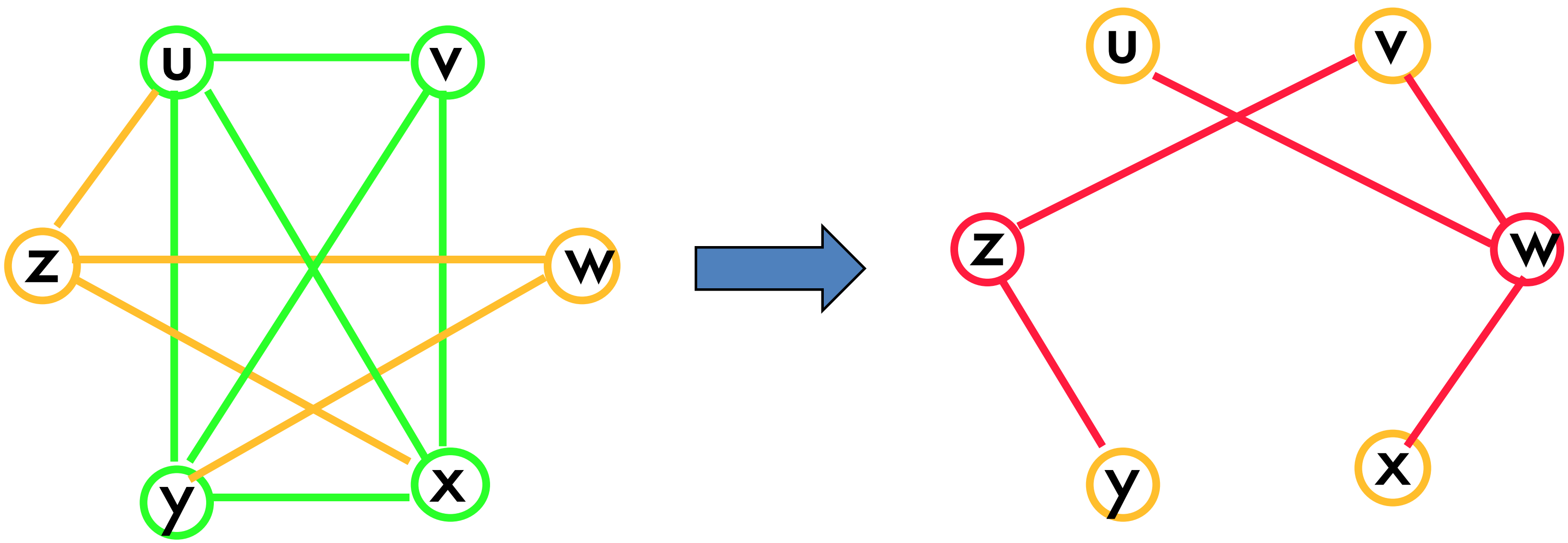
the vertices in V' would only "cover" missing edges and thus are not needed in G_c

CLIQUE vs. VC

- **Theorem**: V' is a clique of G if and only if V – V' is a vertex cover of G_c.

clique
of size 4

VC of
size 2



the vertices in V' would only "cover" missing edges and thus are not needed in G_c

CLIQUE vs. VC

Theorem:V' is a clique of G if and only if V–V' is a vertex cover of G_c.

(=>) G has a clique V'.

Let's assume for the sake of contradiction V-V' is not VC for Gc

Let e'=(a,b) be any edge in Ec that is not covered by V-V' (such edge must exist if V-V' is not VC)

⇒ a not in V-V', and b not in V-V' (by definition of not "covered" for edge (a,b))

⇒ a in V' and b in V', but we also know that V' is a clique in G

⇒ there must be an edge (a,b) in G, and hence e'=(a,b) is in E

If e' is in E, it cannot be in Ec, contradiction.

So V-V' must be a VC for Gc

(<=) V-V' is a VC for Gc.

Let's assume for the sake of contradiction that V' is not a clique in G (i.e. there must be at least one edge missing among V')

⇒ Exist 2 nodes a,b in V' such that edge (a,b) is not in E

⇒ edge (a,b) must be in Ec (be definition of complement)

⇒ but neither a not b are in V-V', so edge (a,b) is not covered by V-V', contradiction with V-V' being a VC for Gc

So V' must be a clique in G

VC and CLIQUE

- We can use previous observation to show that
  - VC $\leq_p$ CLIQUE: given VC(G,k), solve CLIQUE($G'=G_c,k'=|V|-k$)
- And also to show that
  - CLIQUE $\leq_p$ VC: given CLIQUE(G,k), solve VC($G'=G_c,k'=|V|-k$)
- How about IS $\leq_p$ CLIQUE and CLIQUE $\leq_p$ IS?
  - Simple equivalence as well
- Prove that CLIQUE is NPC?
  - CLIQUE is in NP
  - VC $\leq_p$ CLIQUE
  - Direct reduction from 3-SAT quite easy as well [BRV6.4.2]

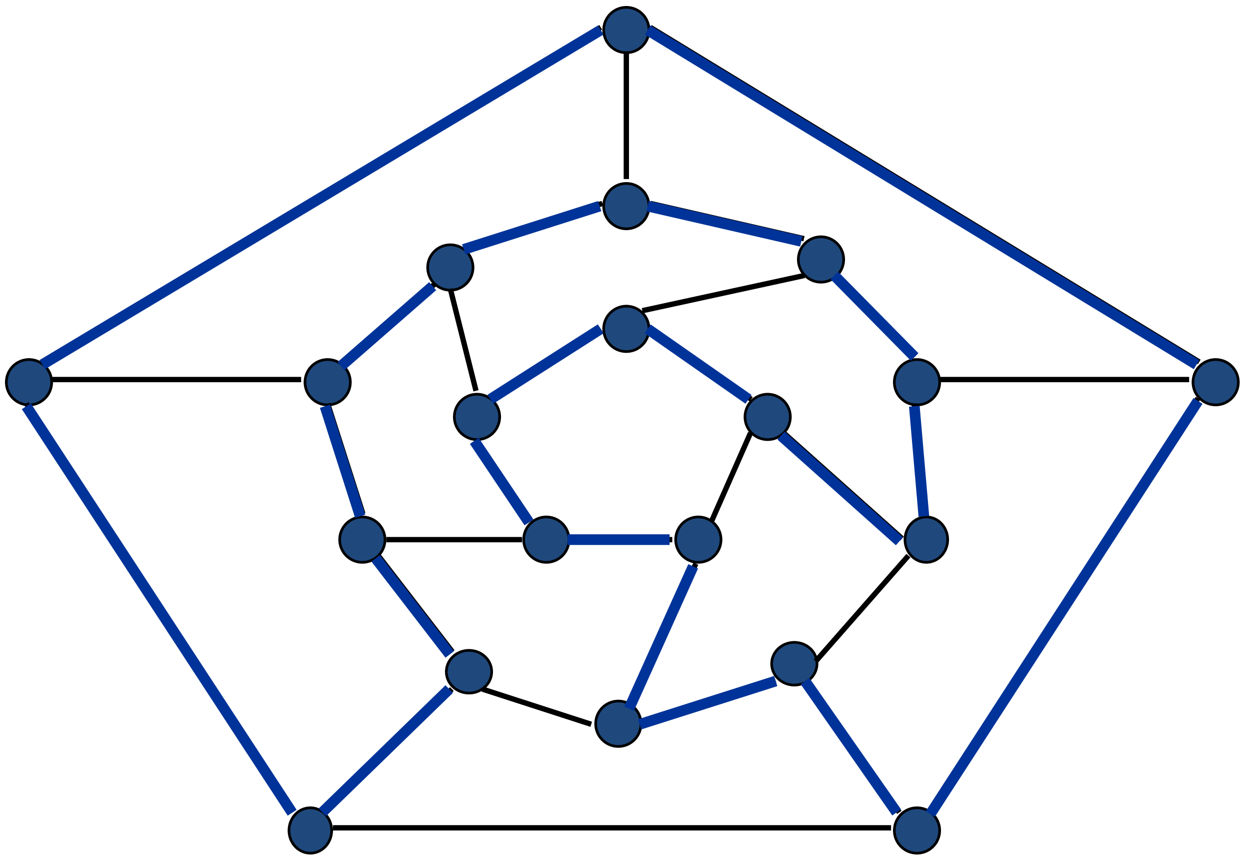
Basic genres.

- Packing problems: SET-PACKING, INDEPENDENT SET.
- Covering problems: SET-COVER, VERTEX-COVER.
- Constraint satisfaction problems: SAT, 3-SAT.
- Sequencing problems: HAMILTONIAN-CYCLE, TSP.
- Partitioning problems: 3D-MATCHING, 3-COLOR.
- Numerical problems: SUBSET-SUM, KNAPSACK.

KT8.5 SEQUENCING PROBLEMS

Hamiltonian Cycle

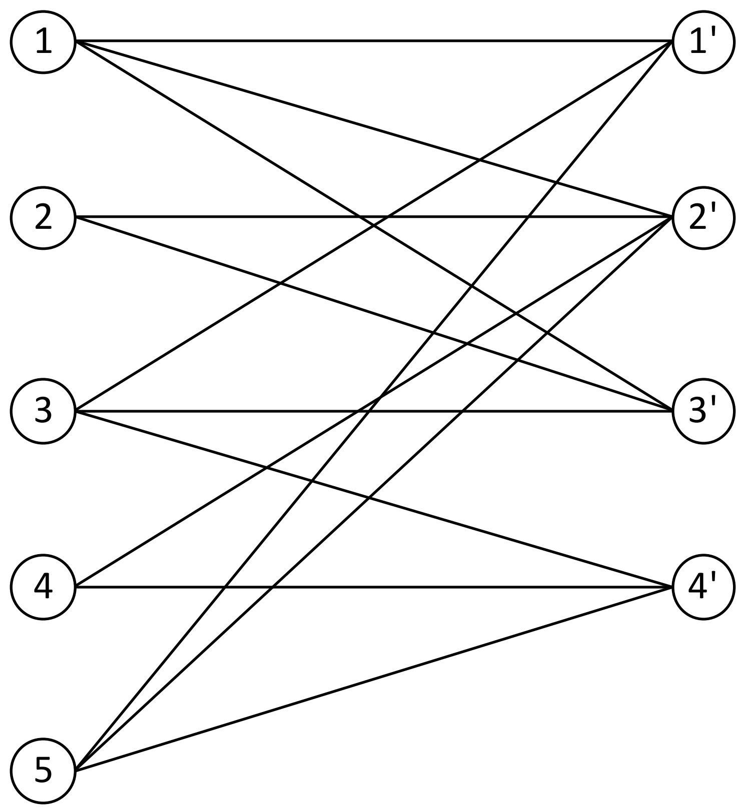
- HAM-CYCLE: given an undirected graph G = (V, E), does there exist a simple cycle Γ that contains every node in V.



YES: vertices and faces of a dodecahedron.

Hamiltonian Cycle

- HAM-CYCLE: given an undirected graph G = (V, E), does there exist a simple cycle Γ that contains every node in V.



NO: bipartite graph with odd number of nodes.

Directed Hamiltonian Cycle

- DIR-HAM-CYCLE: given a digraph G = (V, E), does there exists a simple directed cycle Γ that contains every node in V?

- HAM-CYCLE: given an undirected graph G = (V, E), does there exists a simple directed cycle Γ that contains every node in V?

- DIR-HAM-CYCLE (HAM-CYCLE) is in NP
  - Certificate: Sequence of vertices
  - Certifier: Check if every vertex (except the first) appears exactly once, and that consecutive vertices are connected by a directed (undirected) edge