# Independent sets and vertex covers COMS20010 (Algorithms II)

John Lapinskas, University of Bristol

- Lemmings;
- Pac-Man;
- Minesweeper;
- Tetris;
- Candy Crush;
- Angry Birds;
- Classic Mario games;
- Spelunky;
- Donkey Kong Country 1–3.

- Lemmings;
- Pac-Man;
- Minesweeper;
- Tetris;
- Candy Crush;
- Angry Birds;
- Classic Mario games;
- Spelunky;
- Donkey Kong Country 1-3;

- Lemmings;
- Pac-Man;
- Minesweeper;
- Tetris;
- Candy Crush;
- Angry Birds;
- Classic Mario games;
- Spelunky;
- Donkey Kong Country 1–3;

- Every Legend of Zelda game;
- Every Metroid game;
- Every Fire Emblem game;
- Mainline Pokémon games;
- Mario Kart;
- Desktop Tower Defense;
- Harvest Moon;
- Inventory packing in ARPGs;
- Damage boosting in speedruns.

In a graph G = (V, E), an **independent set** is a subset of V which contains no edges. For example:

In a graph G = (V, E), an **independent set** is a subset of V which contains no edges. For example:



In a graph G = (V, E), an **independent set** is a subset of V which contains no edges. For example:



In a graph G = (V, E), an **independent set** is a subset of V which contains no edges. For example:



In a graph G = (V, E), an **independent set** is a subset of V which contains no edges. For example:



In a graph G = (V, E), an **independent set** is a subset of V which contains no edges. For example:





But this set isn't.

In a graph G = (V, E), an **independent set** is a subset of V which contains no edges. For example:





But this set isn't.

In a graph G = (V, E), an **independent set** is a subset of V which contains no edges. For example:





These sets are independent...

But this set isn't.

Independent sets are important in graphs which model conflicts.

For example, suppose we are trying to assign frequencies to radio transmitters while avoiding interference. If we join two transmitters by an edge when they are close enough to interfere with each other, then we can safely assign the same frequency to all transmitters in an independent set.

Independent sets are important in graphs which model conflicts.

For example, suppose we are trying to assign frequencies to radio transmitters while avoiding interference. If we join two transmitters by an edge when they are close enough to interfere with each other, then we can assign the same frequency to all transmitters in an independent set.

Independent sets are important in graphs which model conflicts.

For example, suppose we are trying to assign frequencies to radio transmitters while avoiding interference. If we join two transmitters by an edge when they are close enough to interfere with each other, then we can assign the same frequency to all transmitters in an independent set.

We would often like to be able to find a **maximum** independent set, i.e. one which is as big as possible.

Independent sets are important in graphs which model conflicts.

For example, suppose we are trying to assign frequencies to radio transmitters while avoiding interference. If we join two transmitters by an edge when they are close enough to interfere with each other, then we can assign the same frequency to all transmitters in an independent set.

We would often like to be able to find a **maximum** independent set, i.e. one which is as big as possible.

The decision version of this problem, **IS**, asks: Given a graph G = (V, E), and an integer k, does G contain an independent set of size at least k?

**Theorem:** IS is NP-complete.

Independent sets are important in graphs which model conflicts.

For example, suppose we are trying to assign frequencies to radio transmitters while avoiding interference. If we join two transmitters by an edge when they are close enough to interfere with each other, then we can assign the same frequency to all transmitters in an independent set.

We would often like to be able to find a **maximum** independent set, i.e. one which is as big as possible.

The decision version of this problem, **IS**, asks: Given a graph G = (V, E), and an integer k, does G contain an independent set of size at least k?

**Theorem:** IS is NP-complete.

We can verify a set is independent in polynomial time, so  $IS \in NP$ .

Independent sets are important in graphs which model conflicts.

For example, suppose we are trying to assign frequencies to radio transmitters while avoiding interference. If we join two transmitters by an edge when they are close enough to interfere with each other, then we can assign the same frequency to all transmitters in an independent set.

We would often like to be able to find a **maximum** independent set, i.e. one which is as big as possible.

The decision version of this problem, **IS**, asks: Given a graph G = (V, E), and an integer k, does G contain an independent set of size at least k?

Theorem: IS is NP-complete.

We can verify a set is independent in polynomial time, so  $IS \in NP$ .

We will show NP-hardness by reducing from 3-SAT, i.e. proving 3-SAT  $\leq_c$  IS. Since we already proved SAT  $\leq_c$  3-SAT, the result follows.

A CNF formula has width 3 if all its OR clauses contain 3 literals. **3-SAT** asks: is the input width-3 CNF formula satisfiable?

**Theorem:** IS is NP-complete. **Goal:** Prove 3-SAT  $\leq_c$  IS.

Let *F* be an instance of 3-SAT. We'll follow our usual approach: build a graph *G* whose size- $(\geq k)$  independent sets correspond to satisfying assignments of *F*, then apply our IS oracle to *G*.

A CNF formula has width 3 if all its OR clauses contain 3 literals. **3-SAT** asks: is the input width-3 CNF formula satisfiable?

**Theorem:** IS is NP-complete. **Goal:** Prove 3-SAT  $\leq_c$  IS.

Let *F* be an instance of 3-SAT. We'll follow our usual approach: build a graph *G* whose size- $(\geq k)$  independent sets correspond to satisfying assignments of *F*, then apply our IS oracle to *G*.

To simulate the variables of F, we want a gadget that can be in one of two states which will represent True and False...

A CNF formula has width 3 if all its OR clauses contain 3 literals. **3-SAT** asks: is the input width-3 CNF formula satisfiable?

**Theorem:** IS is NP-complete. **Goal:** Prove 3-SAT  $\leq_c$  IS.

Let *F* be an instance of 3-SAT. We'll follow our usual approach: build a graph *G* whose size- $(\geq k)$  independent sets correspond to satisfying assignments of *F*, then apply our IS oracle to *G*.

To simulate the variables of F, we want a gadget that can be in one of two states which will represent True and False...

$$x \longrightarrow$$
 False  $x$  True

Idea: Use an edge!

A CNF formula has width 3 if all its OR clauses contain 3 literals. **3-SAT** asks: is the input width-3 CNF formula satisfiable?

**Theorem:** IS is NP-complete. **Goal:** Prove 3-SAT  $\leq_c$  IS.

Let *F* be an instance of 3-SAT. We'll follow our usual approach: build a graph *G* whose size- $(\geq k)$  independent sets correspond to satisfying assignments of *F*, then apply our IS oracle to *G*.

To simulate the variables of F, we want a gadget that can be in one of two states which will represent True and False...

$$x \longrightarrow$$
 False  $\longrightarrow$  True

Idea: Use an edge!

An independent set can't contain both vertices, and (if we do everything else right) a **maximum** independent set must contain one of the two vertices.

A CNF formula has width 3 if all its OR clauses contain 3 literals. **3-SAT** asks: is the input width-3 CNF formula satisfiable?

**Theorem:** IS is NP-complete. **Goal:** Prove 3-SAT  $\leq_c$  IS.

Let F be an instance of 3-SAT. We'll build a graph G whose size- $(\geq k)$  independent sets correspond to satisfying assignments of F.

Our variable gadget is:

$$x \longrightarrow$$
 False  $x$  True

We use the same idea to model the clauses of F. We have three literals in the clause, and we want to force one of them to be true, so...

A CNF formula has width 3 if all its OR clauses contain 3 literals. **3-SAT** asks: is the input width-3 CNF formula satisfiable?

**Theorem:** IS is NP-complete. **Goal:** Prove 3-SAT  $\leq_c$  IS.

Let F be an instance of 3-SAT. We'll build a graph G whose size- $(\geq k)$  independent sets correspond to satisfying assignments of F.

Our variable gadget is:

$$x \longrightarrow False \bigoplus x \bigoplus True$$

We use the same idea to model the clauses of F. We have three literals in the clause, and we want to force one of them to be true, so...

\_\_\_\_\_

$$x \lor \neg y \lor z \longrightarrow x$$
 True  $z$  True

We will set things up so that:

Maximum independent set  $\implies$  exactly one vertex is included.

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G?

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G?

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build *G*? *w* False *w* True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build *G*? *w* False *w* True *x* False *x* True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build *G*? *w* False *w* True *x* False *x* True *y* False *y* True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G?

w False w True x False x True y False y True z False z True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G?

w False w True x False x True y False y True z False z True



We add gadgets for each variable... and gadgets for each clause...

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True





We add gadgets for each variable... and gadgets for each clause...

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



We add gadgets for each variable... and gadgets for each clause...

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build *G*? *w* False *w* True *x* False *x* True *y* False *y* True *z* False *z* True



And now we need to enforce that "x True" is only in the independent set if x is actually true. Which we can do by joining it to the corresponding "x False" vertex in the variable gadget.

| John L | apinskas. |
|--------|-----------|
|--------|-----------|

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



And now we need to enforce that "x True" is only in the independent set if x is actually true. Which we can do by joining it to the corresponding "x False" vertex in the variable gadget.

| John La | apinskas |
|---------|----------|
|---------|----------|

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



And now we need to enforce that "x True" is only in the independent set if x is actually true. Which we can do by joining it to the corresponding "x False" vertex in the variable gadget.
So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build *G*? *w* False *w* True *x* False *x* True *y* False *y* True *z* False *z* True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build *G*? *w* False *w* True *x* False *x* True *y* False *y* True *z* False *z* True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build *G*? *w* False *w* True *x* False *x* True *y* False *y* True *z* False *z* True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



Now a satisfying assignment in F, say w = True, x = True, y = False, z = False, corresponds to a size-6 independent set in G.

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



Conversely, any independent set in G has at most one vertex from each variable gadget...

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G?

w False w True x False x True y False y True z False z True



Conversely, any independent set in G has at most one vertex from each variable gadget...

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G?

w False w True x False x True y False y True z False z True



Conversely, any independent set in G has at most one vertex from each of these edges... and at most one vertex from each clause gadget, for a total of at most 6 vertices.

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G?

w False w True x False x True y False y True z False z True



Conversely, any independent set in G has at most one vertex from each of these edges... and at most one vertex from each clause gadget, for a total of at most 6 vertices.

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G?

w False w True x False x True y False y True z False z True



So any size- $(\geq 6)$  independent set must have size **exactly** 6, and contain one vertex from each variable gadget and one from each clause gadget.

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



So any size- $(\geq 6)$  independent set must have size **exactly** 6, and contain one vertex from each variable gadget and one from each clause gadget.

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



Hence any size-( $\geq 6$ ) independent set corresponds to an assignment, in this case w = x = y = False, z = True. It must be satisfying because there are no edges between variable vertices and clause vertices.

So if  $F = (x \lor \neg y \lor z) \land (w \lor \neg x \lor \neg z)$ , say, how do we build G? w False w True x False x True y False y True z False z True



The same construction (and the same correctness proof) works for any instance of 3-SAT.  $\hfill\square$ 

#### Recall from Video 8-2...

A vertex cover in a graph G = (V, E) is a set  $X \subseteq V$  such that every edge in E has at least one vertex in X.







Not a valid vertex cover.

## Recall from Video 8-2...

A vertex cover in a graph G = (V, E) is a set  $X \subseteq V$  such that every edge in E has at least one vertex in X.



A valid vertex cover.

Not a valid vertex cover.

We would like to find the **smallest possible** vertex cover of G. We claimed this problem was hard to solve exactly (our algorithm was approximate), but we never proved it...

# Recall from Video 8-2...

A vertex cover in a graph G = (V, E) is a set  $X \subseteq V$  such that every edge in E has at least one vertex in X.



A valid vertex cover.

Not a valid vertex cover.

We would like to find the **smallest possible** vertex cover of G. We claimed this problem was hard to solve exactly (our algorithm was approximate), but we never proved it...

The decision version of this problem (VC) asks: Given a graph G, and an integer k, does G contain a vertex cover of size at most k?

VC asks: Does G contain a vertex cover of size at most k?

VC asks: Does G contain a vertex cover of size at most k?

Theorem: VC is NP-complete.

VC asks: Does G contain a vertex cover of size at most k?

#### Theorem: VC is NP-complete.

We can verify a set is a vertex cover in polynomial time, so  $VC \in NP$ . We'll prove NP-hardness by proving IS  $\leq_c VC$ .

VC asks: Does G contain a vertex cover of size at most k?

#### Theorem: VC is NP-complete.

We can verify a set is a vertex cover in polynomial time, so  $VC \in NP$ . We'll prove NP-hardness by proving IS  $\leq_c VC$ .

This time though, we'll do it non-constructively, without gadgets.

VC asks: Does G contain a vertex cover of size at most k?

Theorem: VC is NP-complete.

We can verify a set is a vertex cover in polynomial time, so  $VC \in NP$ . We'll prove NP-hardness by proving IS  $\leq_c VC$ .

This time though, we'll do it **non-constructively**, without gadgets.

**Lemma:** X is an independent set if and only if  $V \setminus X$  is a vertex cover.

**VC** asks: Does G contain a vertex cover of size at most k?

Theorem: VC is NP-complete.

We can verify a set is a vertex cover in polynomial time, so  $VC \in NP$ . We'll prove NP-hardness by proving IS  $\leq_c VC$ .

This time though, we'll do it **non-constructively**, without gadgets.

**Lemma:** X is an independent set if and only if  $V \setminus X$  is a vertex cover. (Because an edge intersects  $V \setminus X$  if and only if it's **not** a subset of X.)

VC asks: Does G contain a vertex cover of size at most k?

Theorem: VC is NP-complete.

We can verify a set is a vertex cover in polynomial time, so  $VC \in NP$ . We'll prove NP-hardness by proving IS  $\leq_c VC$ .

This time though, we'll do it **non-constructively**, without gadgets.

**Lemma:** X is an independent set if and only if  $V \setminus X$  is a vertex cover. (Because an edge intersects  $V \setminus X$  if and only if it's **not** a subset of X.)

So G contains an independent set of size at **least** k if and only if G contains a vertex cover of size at **most** |V| - k.

VC asks: Does G contain a vertex cover of size at most k?

**Theorem:** VC is NP-complete.

We can verify a set is a vertex cover in polynomial time, so  $VC \in NP$ . We'll prove NP-hardness by proving IS  $\leq_c VC$ .

This time though, we'll do it non-constructively, without gadgets.

**Lemma:** X is an independent set if and only if  $V \setminus X$  is a vertex cover. (Because an edge intersects  $V \setminus X$  if and only if it's **not** a subset of X.)

So G contains an independent set of size at **least** k if and only if G contains a vertex cover of size at **most** |V| - k.

Our reduction just passes the instance (G, |V| - k) to our VC-oracle.

VC asks: Does G contain a vertex cover of size at most k?

Theorem: VC is NP-complete.

VC asks: Does G contain a vertex cover of size at most k?

**Theorem:** VC is NP-complete.

In video 8-2, we expressed finding maximum vertex covers in terms of **integer** linear programming for our approximation algorithm:



 $\sum_{v} x_{v} \rightarrow \text{min subject to}$   $x_{u} + x_{v} \geq 1 \text{ for all } \{u, v\} \in E;$   $x_{v} \leq 1 \text{ for all } v \in V;$   $x_{v} \geq 0 \text{ for all } v \in V;$   $x_{v} \in \mathbb{N} \text{ for all } v \in V.$ 

VC asks: Does G contain a vertex cover of size at most k?

**Theorem:** VC is NP-complete.

In video 8-2, we expressed finding maximum vertex covers in terms of **integer** linear programming for our approximation algorithm:



 $\sum_{v} x_{v} \rightarrow \text{min subject to}$   $x_{u} + x_{v} \geq 1 \text{ for all } \{u, v\} \in E;$   $x_{v} \leq 1 \text{ for all } v \in V;$   $x_{v} \geq 0 \text{ for all } v \in V;$   $x_{v} \in \mathbb{N} \text{ for all } v \in V.$ 

**Corollary:** Integer linear programming is NP-hard!

VC asks: Does G contain a vertex cover of size at most k?

**Theorem:** VC is NP-complete.

In video 8-2, we expressed finding maximum vertex covers in terms of **integer** linear programming for our approximation algorithm:



 $\sum_{v} x_{v} \rightarrow \text{min subject to}$   $x_{u} + x_{v} \geq 1 \text{ for all } \{u, v\} \in E;$   $x_{v} \leq 1 \text{ for all } v \in V;$   $x_{v} \geq 0 \text{ for all } v \in V;$   $x_{v} \in \mathbb{N} \text{ for all } v \in V.$ 

**Corollary:** Integer linear programming is NP-hard!

Notice we reduced SAT  $\leq_c 3$ -SAT  $\leq_c IS \leq_c VC \leq_c ILP$  — by proving one problem is NP-hard, we make all our future hardness proofs easier...