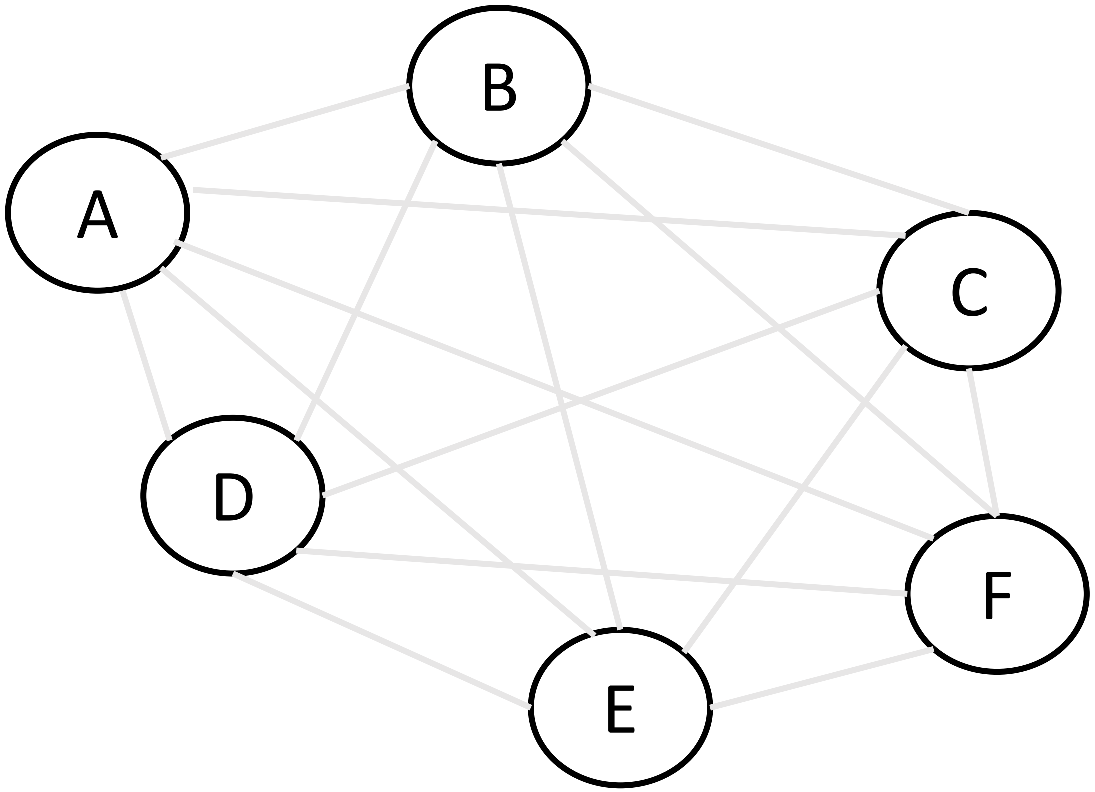
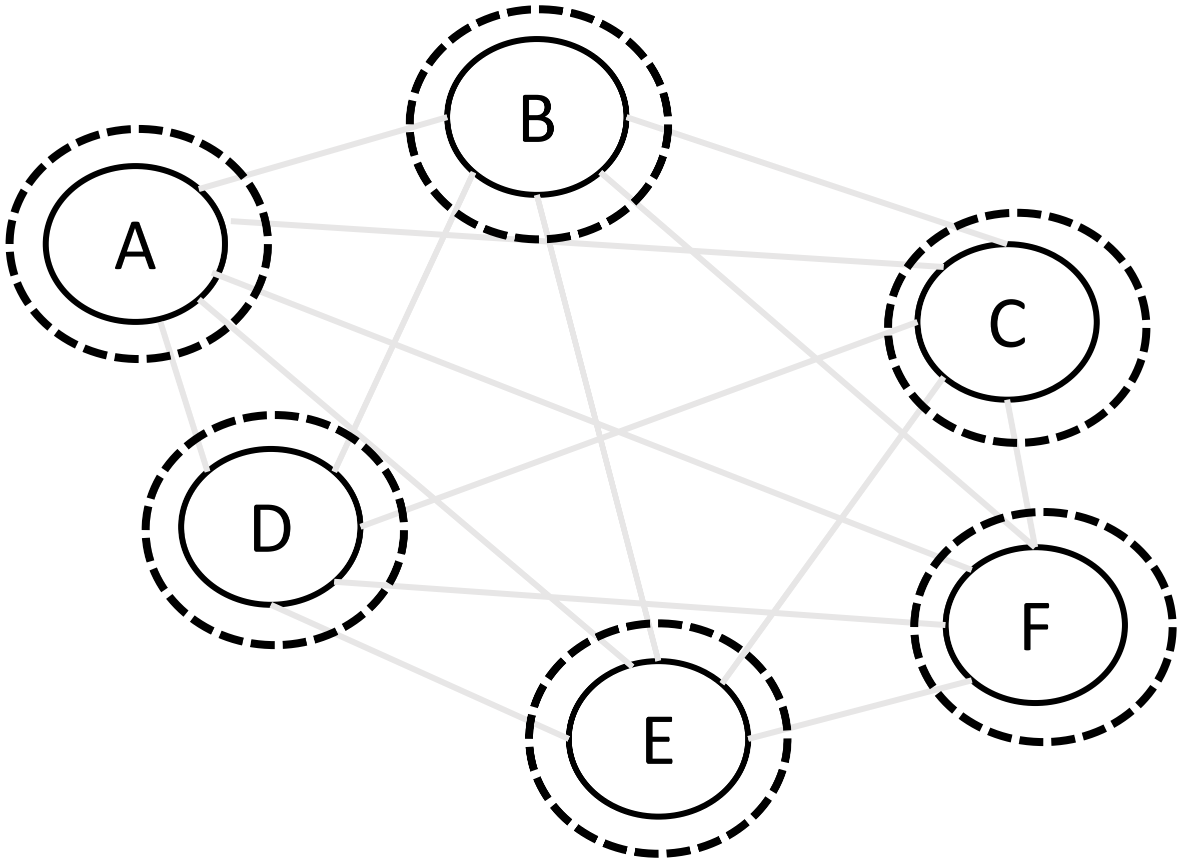
We imagined an instance of problem…



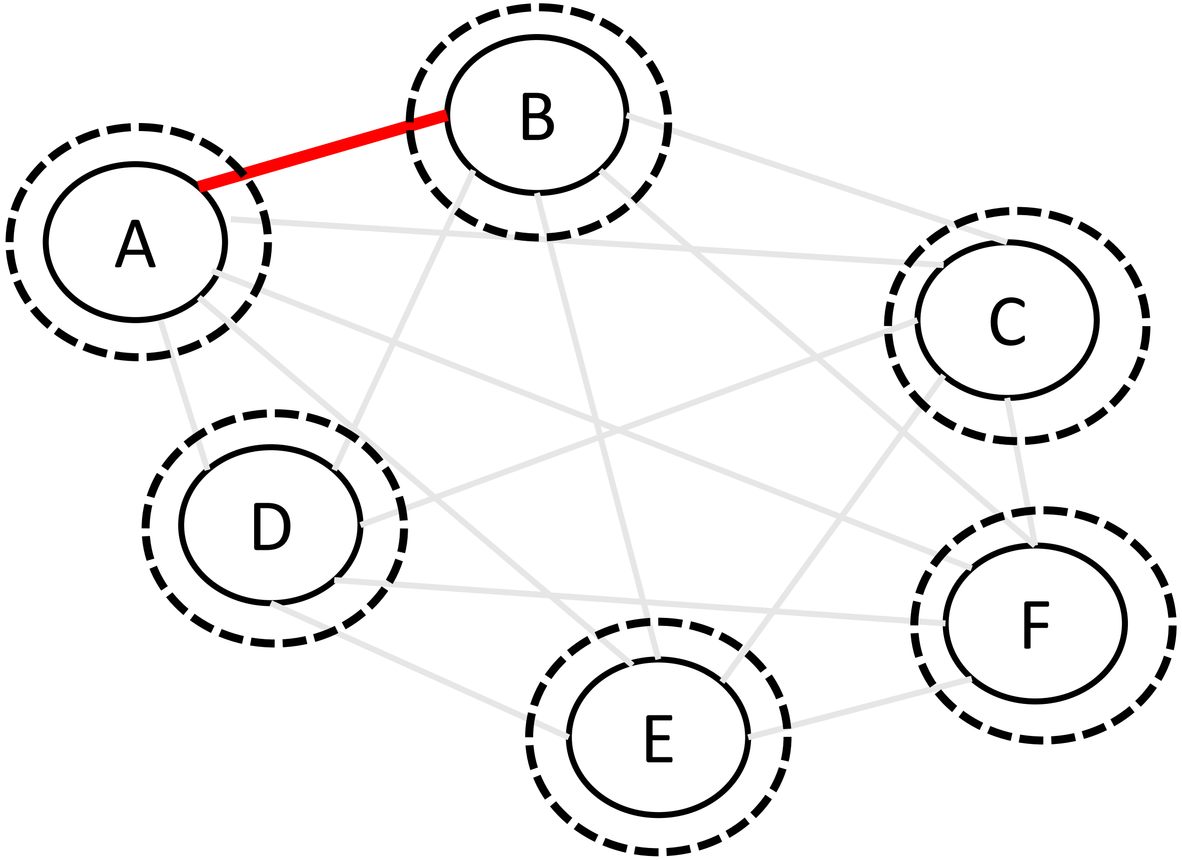
$k = 3$

And reasoned about a solution $\mathcal{G}$ made by our algorithm:



$k = 3$

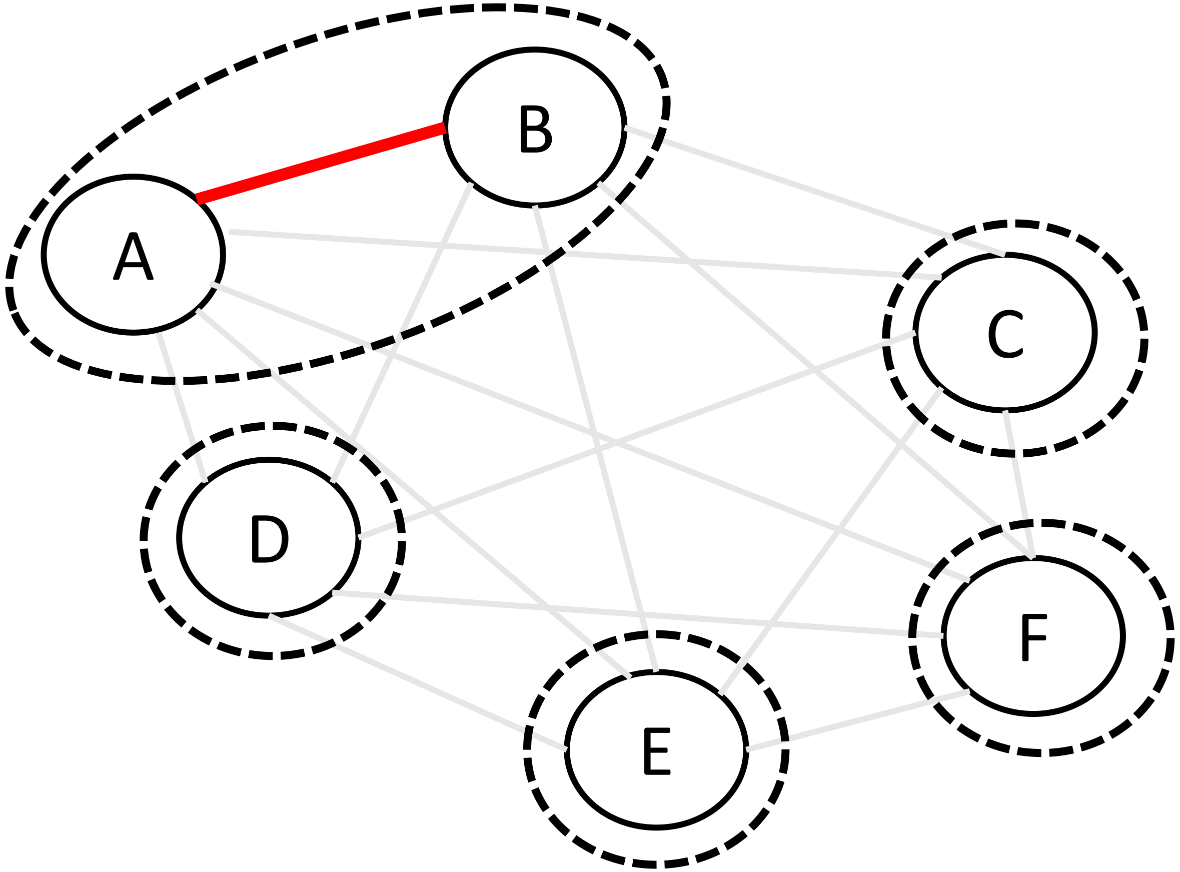
And reasoned about a solution $\mathcal{G}$ made by our algorithm:



$k = 3$

Red edges = E' edges that merged in Step 4

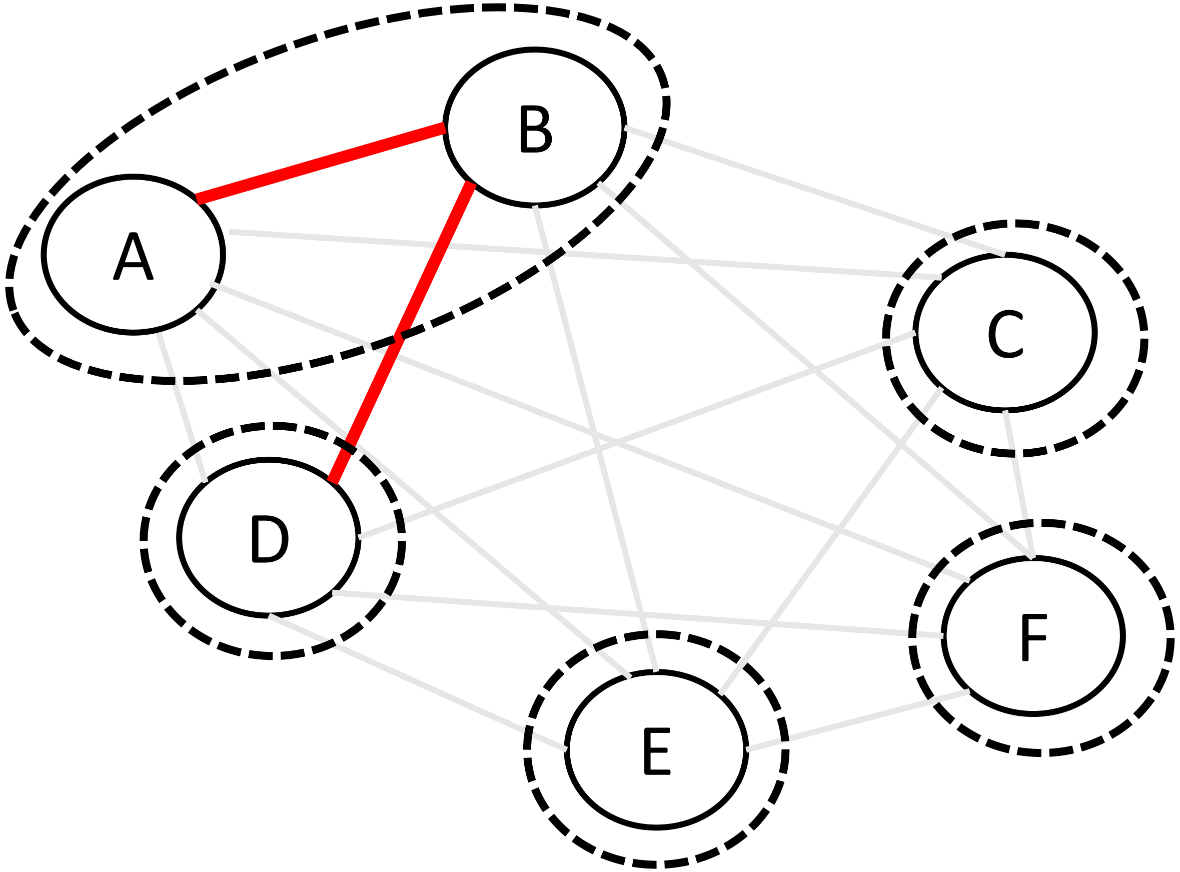
And reasoned about a solution $\mathcal{G}$ made by our algorithm:



$k = 3$

Red edges = E' edges that merged in Step 4

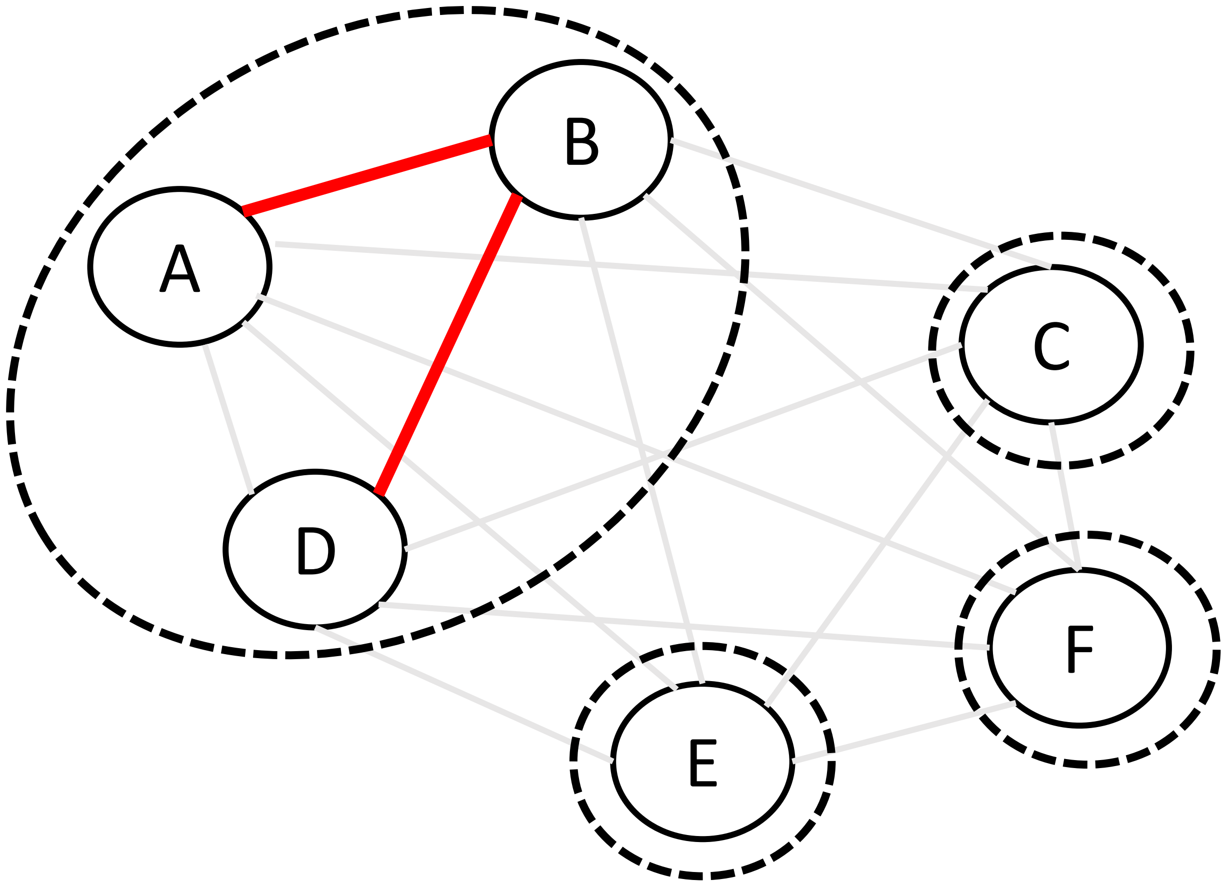
And reasoned about a solution $\mathcal{G}$ made by our algorithm:



$k = 3$

Red edges = E' edges that merged in Step 4

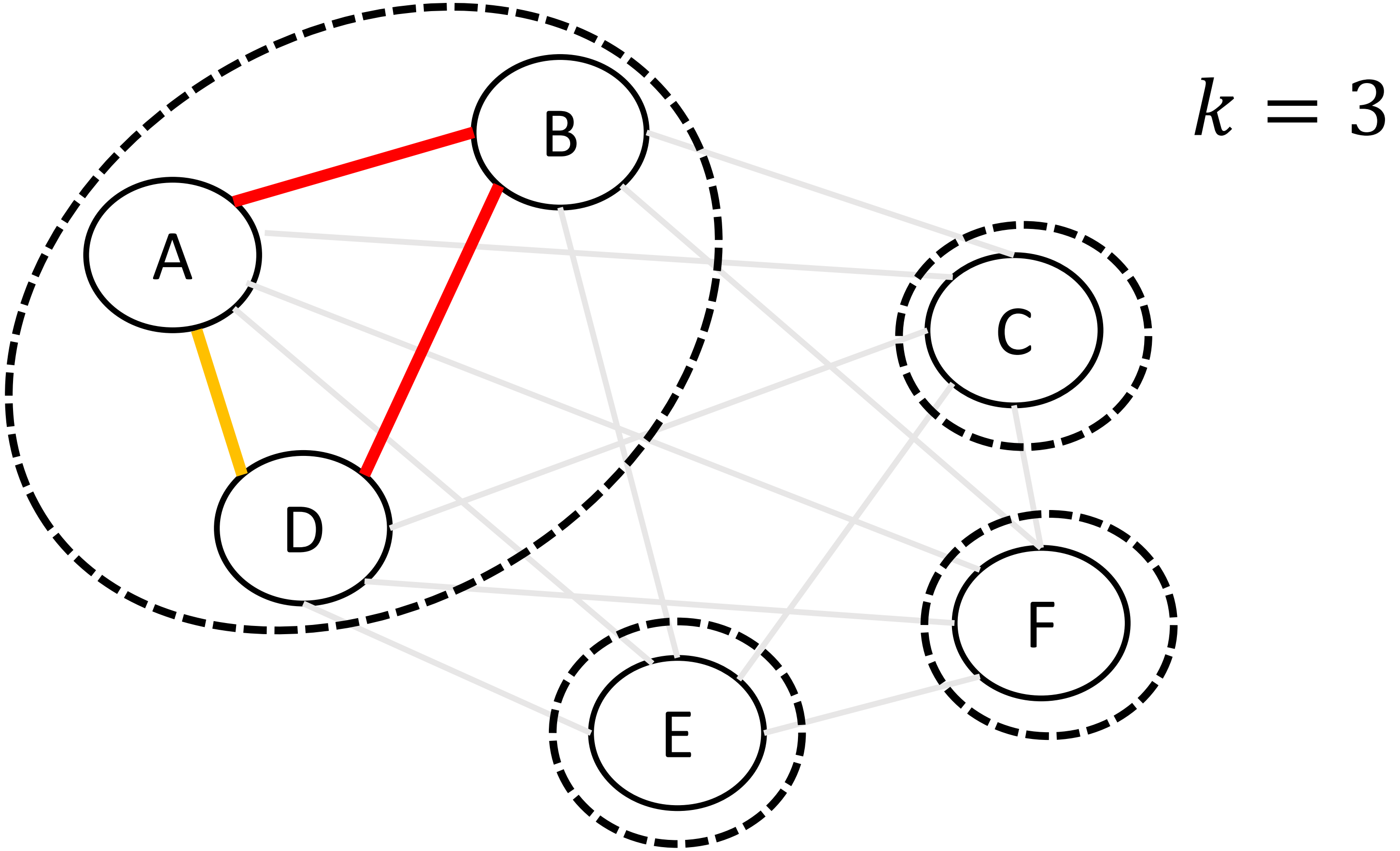
And reasoned about a solution $\mathcal{G}$ made by our algorithm:



$k = 3$

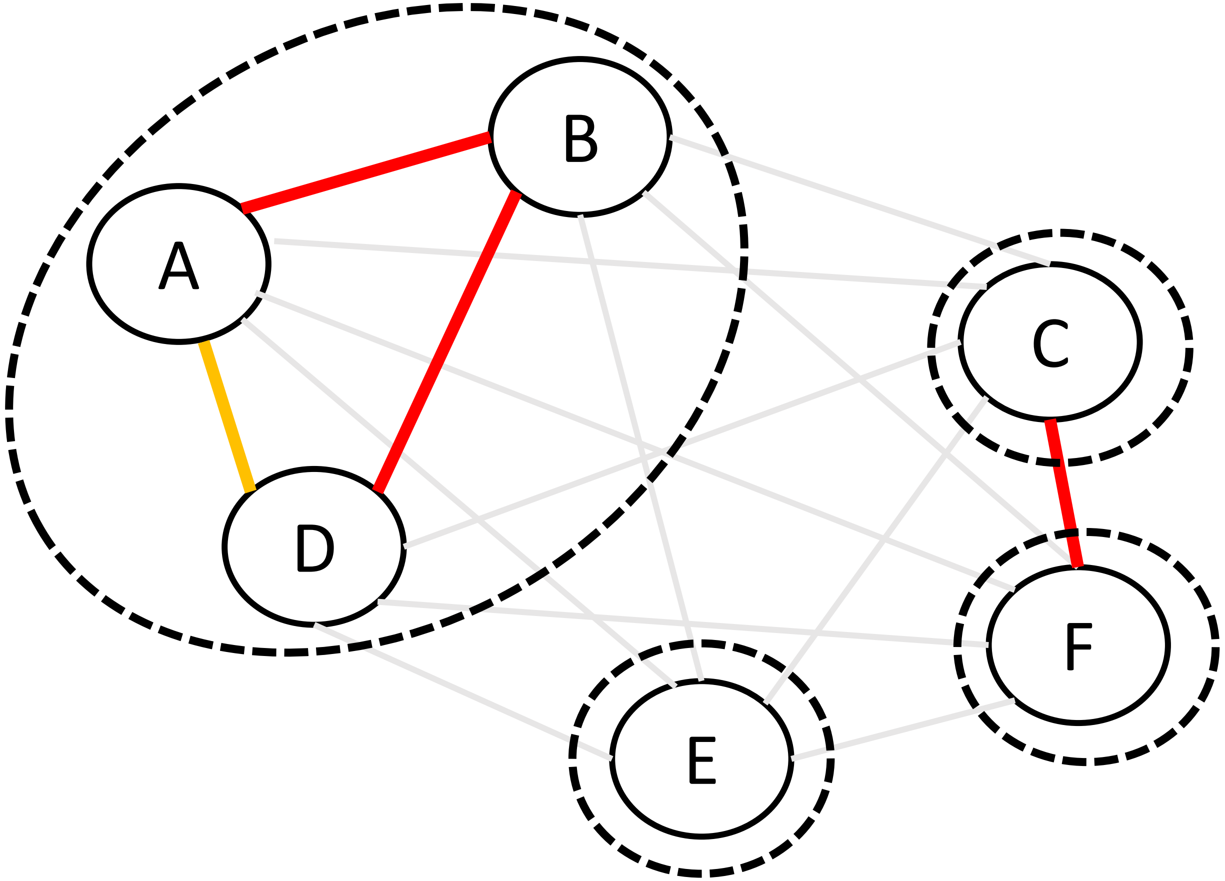
Red edges = E' edges that merged in Step 4

And reasoned about a solution $\mathcal{G}$ made by our algorithm:



Red edges = E' edges that merged in Step 4
Orange edges = E' edges that became intra-category "for free"

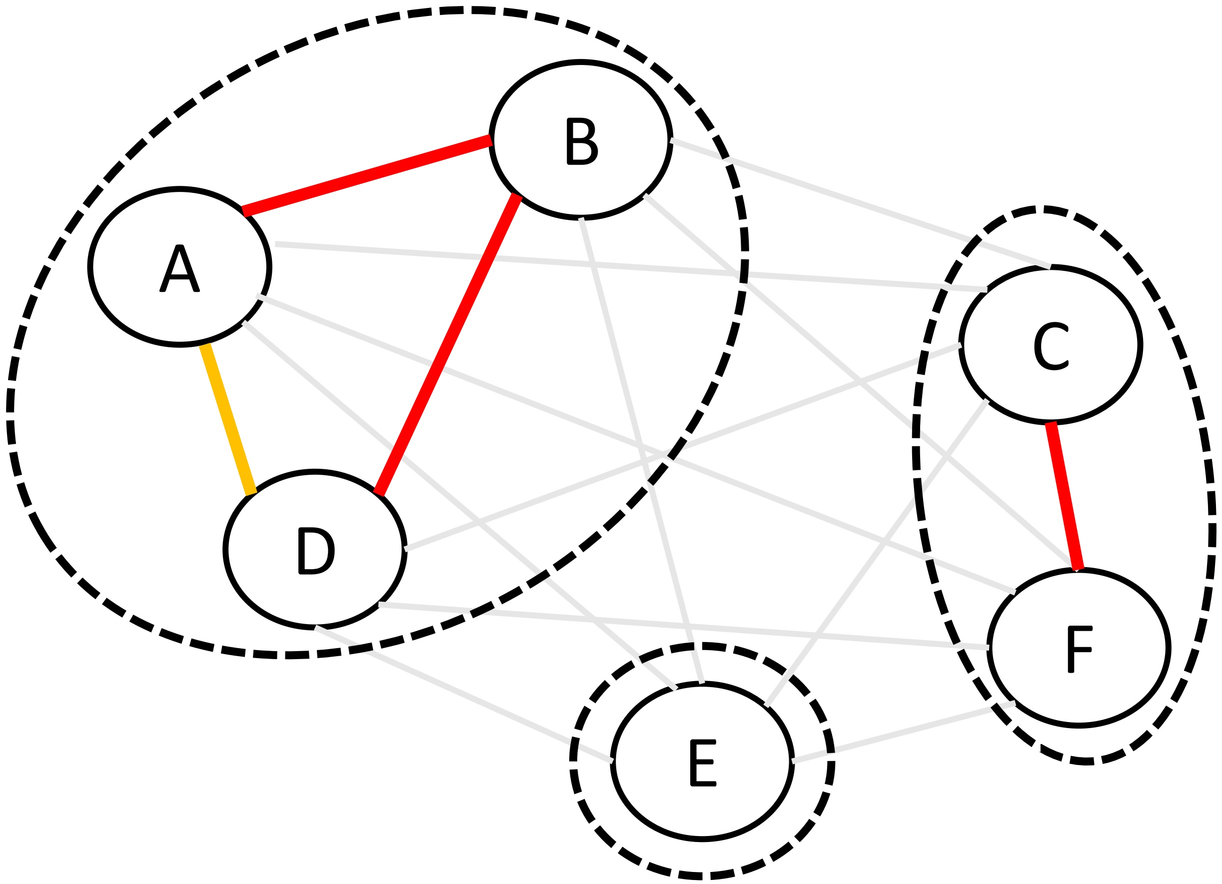
And reasoned about a solution $\mathcal{G}$ made by our algorithm:



$k = 3$

Red edges = E' edges that merged in Step 4
Orange edges = E' edges that became intra-category for free

And reasoned about a solution $\mathcal{G}$ made by our algorithm:

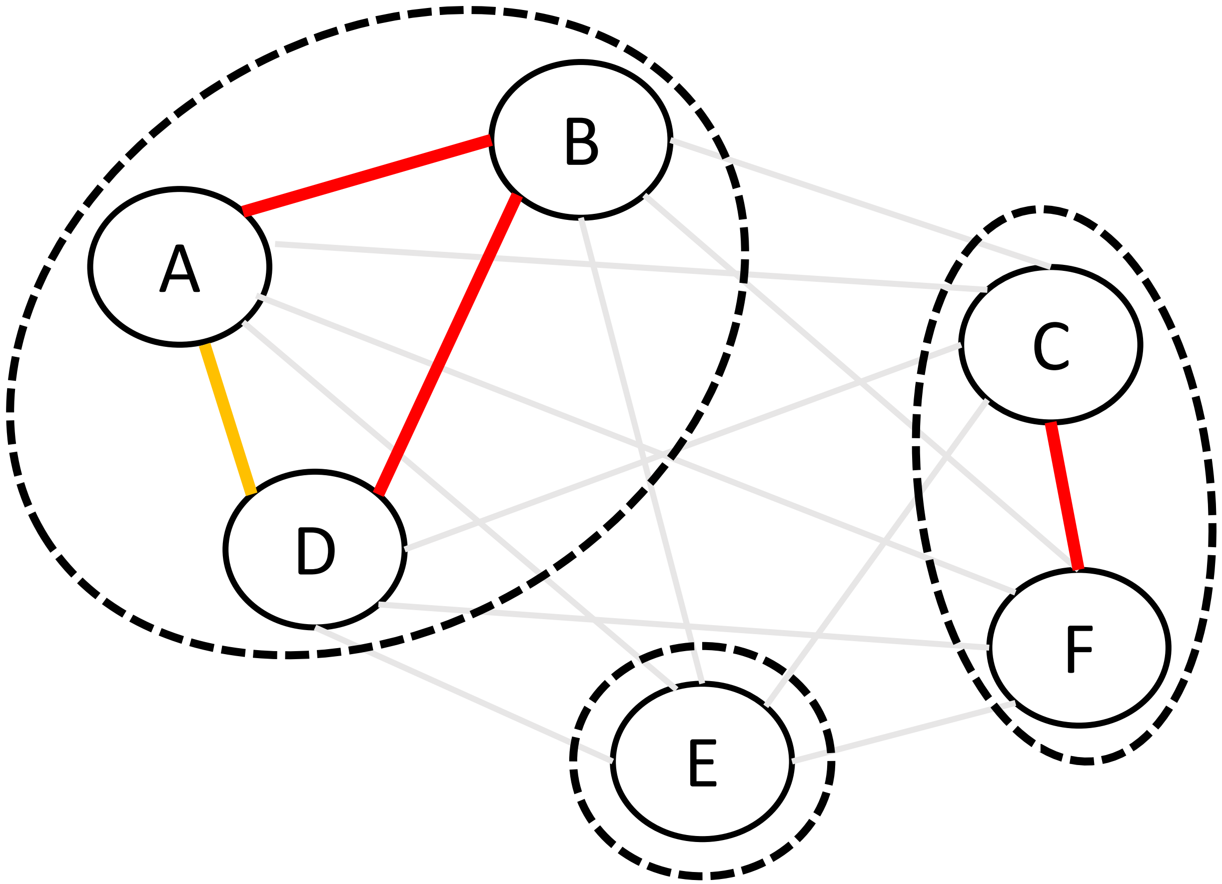


$k = 3$

Red edges = E' edges that merged in Step 4

Orange edges = E' edges that became intra-category for free

And reasoned about a solution $\mathcal{G}$ made by our algorithm:

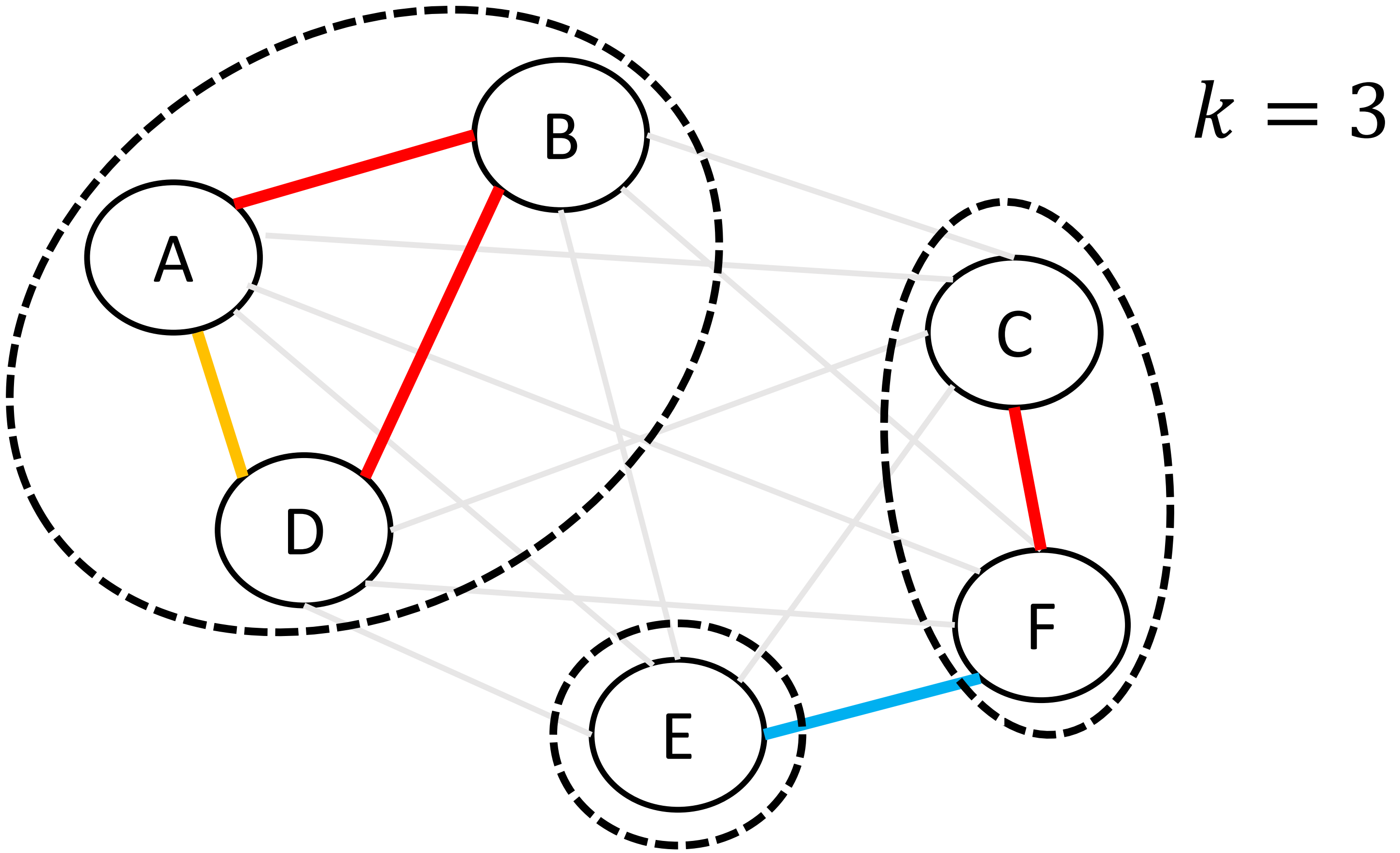


$k = 3$

Done! A complete greedy solution

Red edges = E' edges that merged in Step 4
Orange edges = E' that became intra-category for free in Step 4

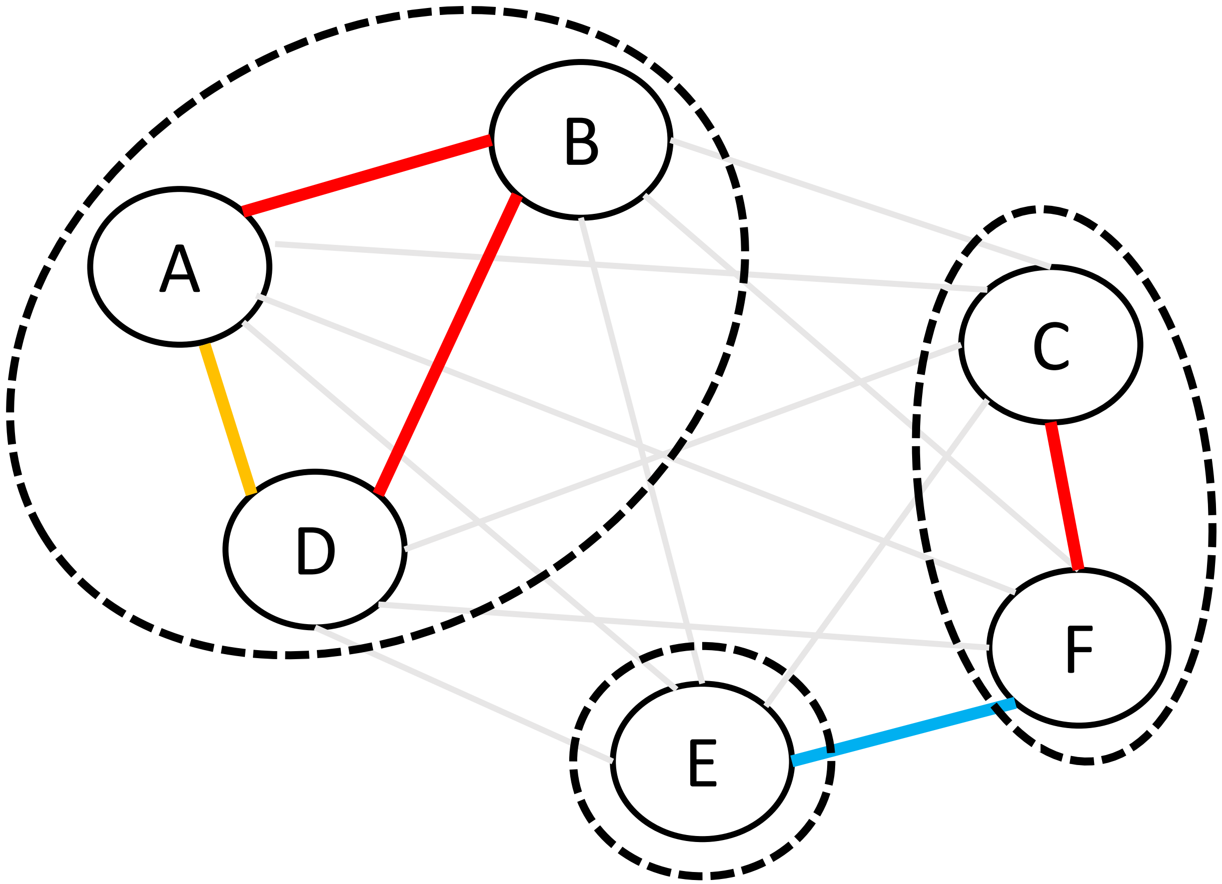
And reasoned about a solution $\mathcal{G}$ made by our algorithm:



red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

And reasoned about a solution $\mathcal{G}$ made by our algorithm:

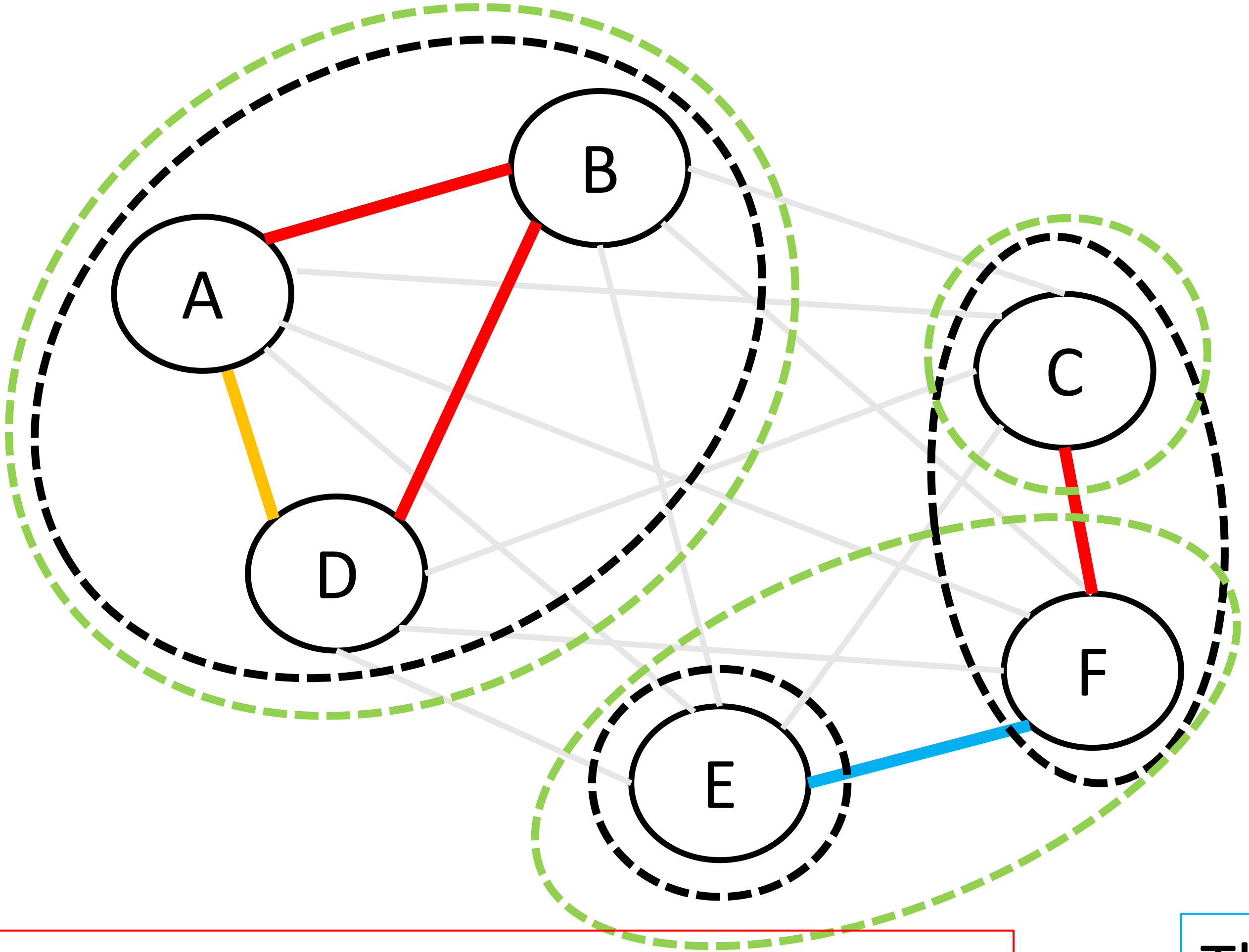


Proceeding with the proof: compare $\mathcal{G}$ to a **different** optimal solution $\mathcal{O}$

red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

Case 1: some edges of E' are inter-category in $\mathcal{O}$

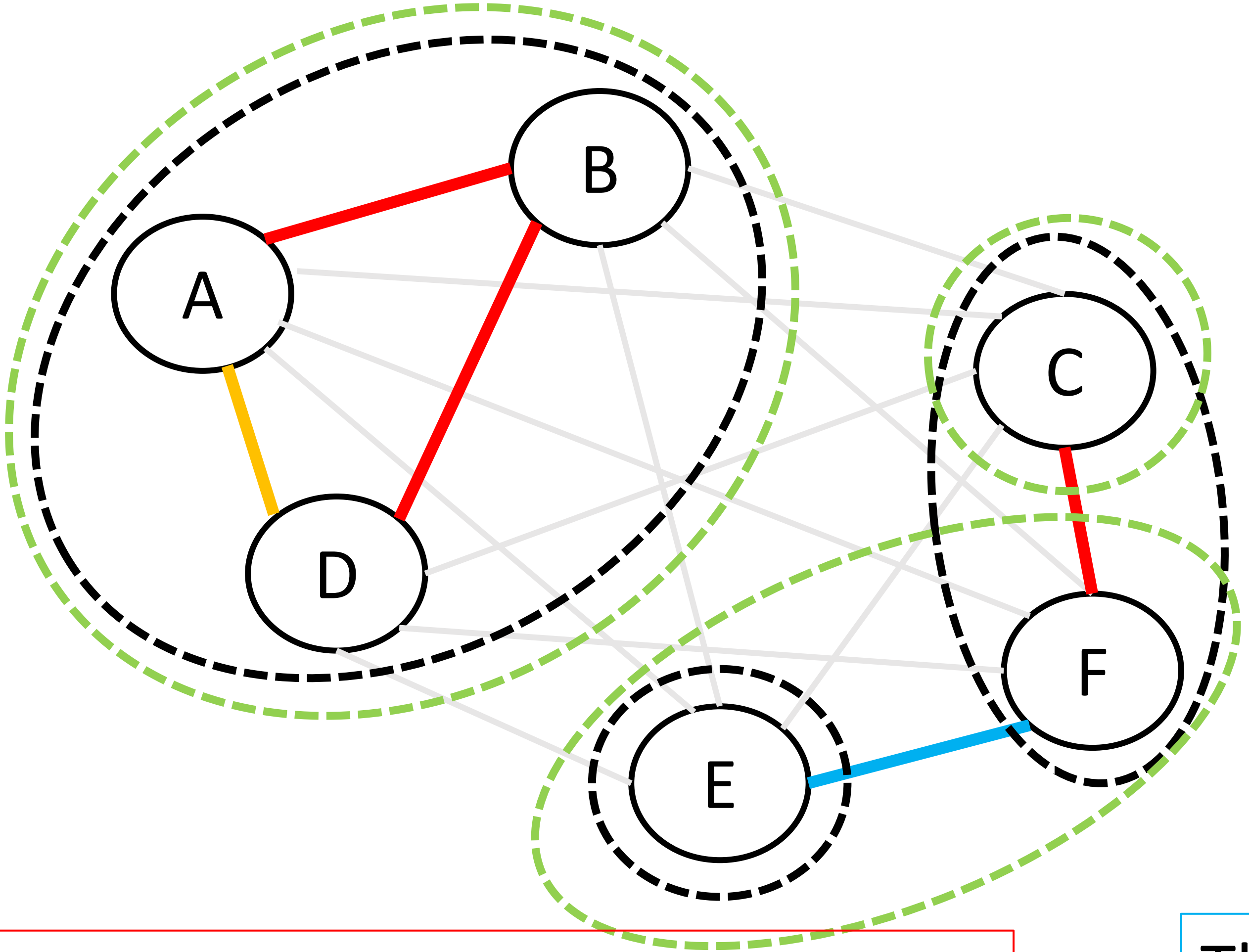


Example: $\mathcal{O}$ is in green.

red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

Case 1: some edges of E' are inter-category in $\mathcal{O}$



Example: $\mathcal{O}$ is in green.

Which edge defines Cost($\mathcal{O}$)?

A. (C,F)

B. (E,F)

C. Another red edge

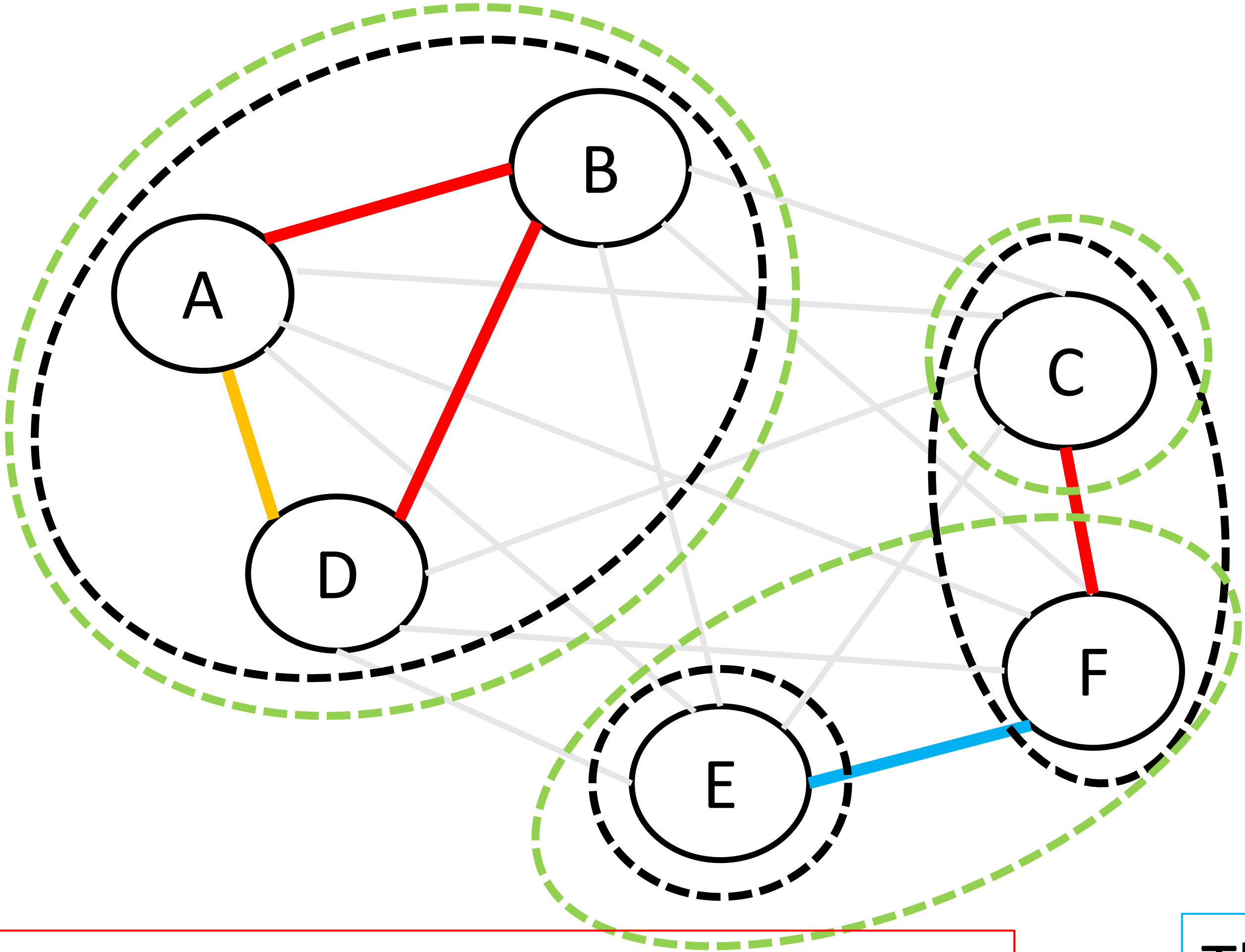
D. Impossible to determine

red edges = E' edges merged in Step 4

Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

Case 1: some edges of E' are inter-category in $\mathcal{O}$



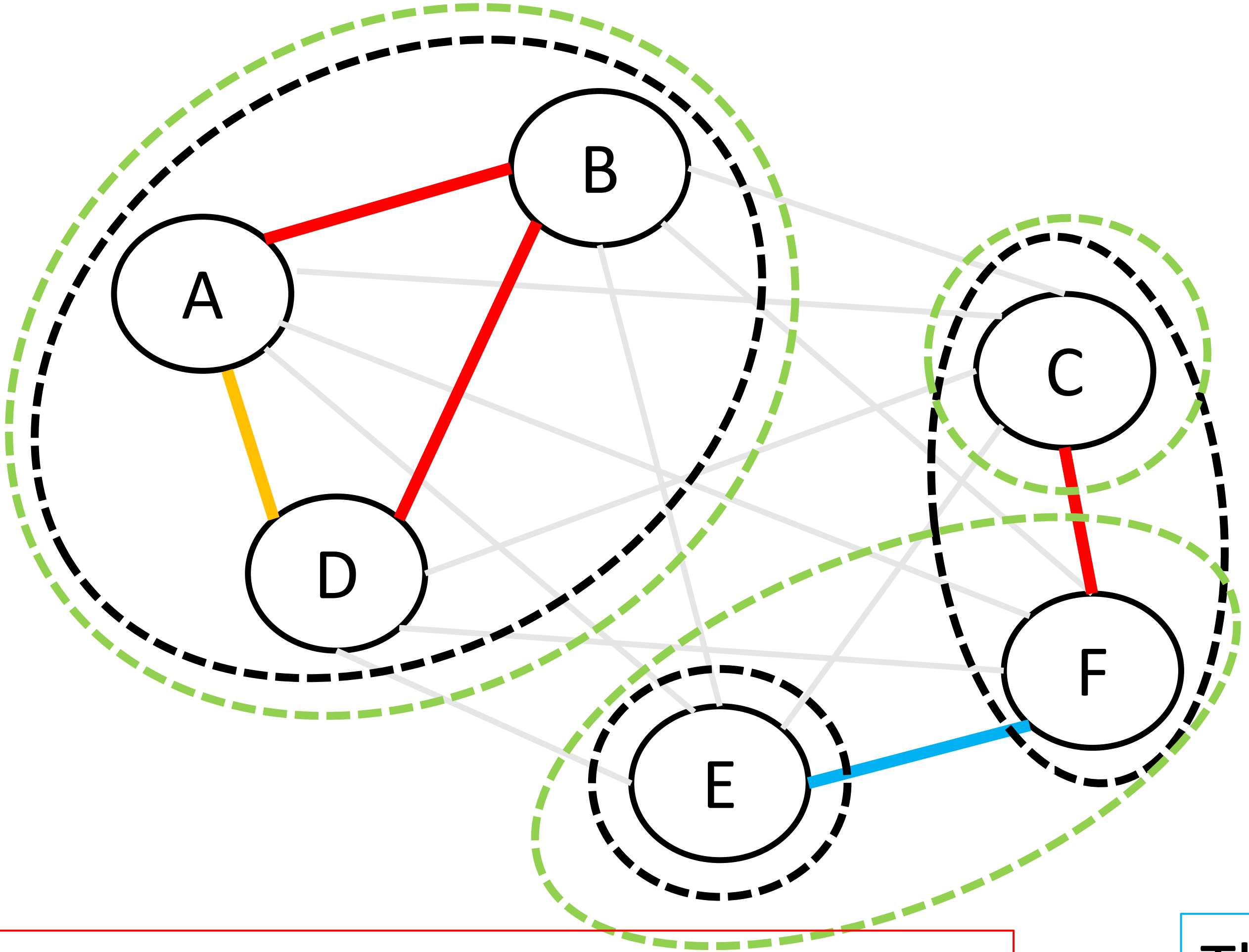
Example: $\mathcal{O}$ is in green.
(C,F) defines Cost($\mathcal{O}$). How does the weight of (C,F) compare to the weight (E,F)?

A. weight(C,F) $\leq$ weight(E,F)
B. weight(C,F) $\geq$ weight(E,F)
C. Impossible to determine

red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

Case 1: some edges of E' are inter-category in $\mathcal{O}$

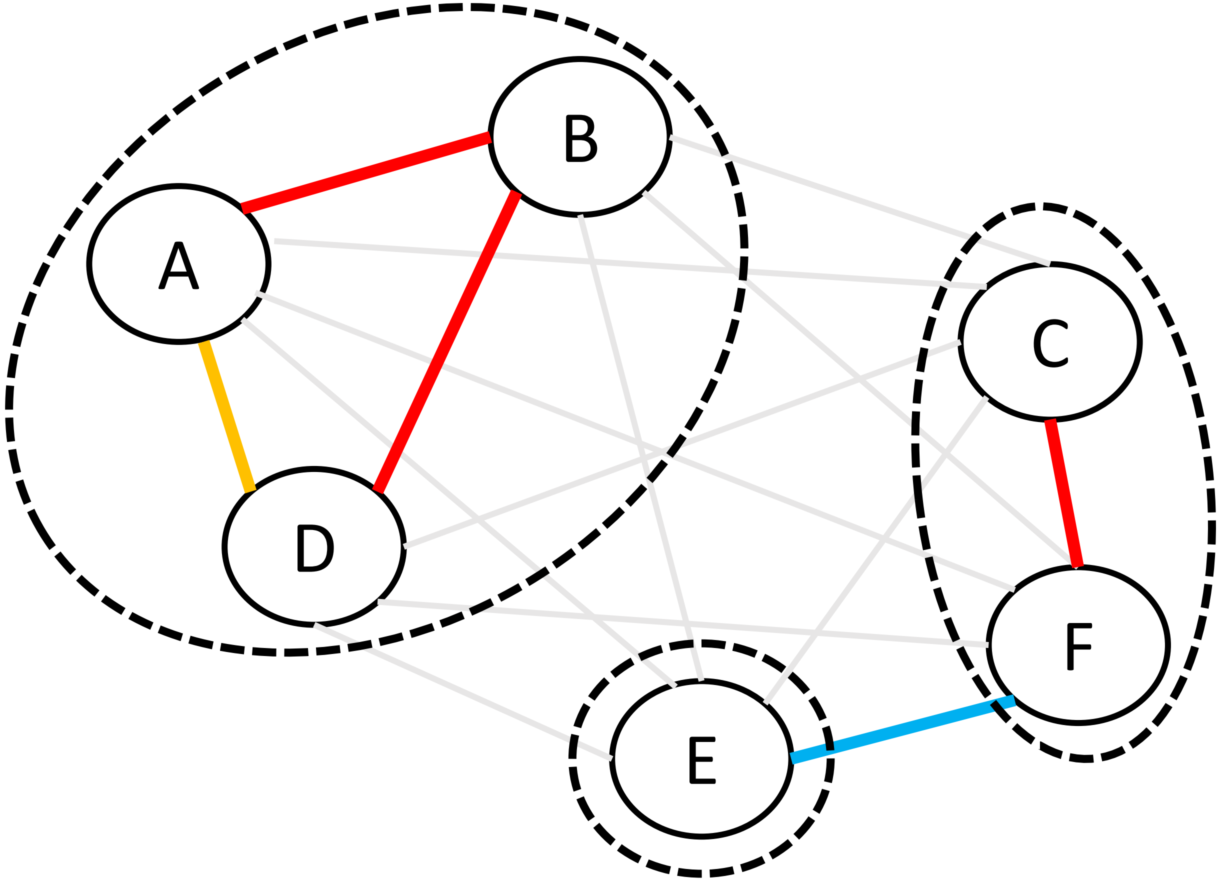


Example: $\mathcal{O}$ is in green.
(C,F) defines Cost($\mathcal{O}$), and has **equal or greater weight** than the blue edge:
$\rightarrow$ Cost($\mathcal{G}$) $\leq$ Cost($\mathcal{O}$)

red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

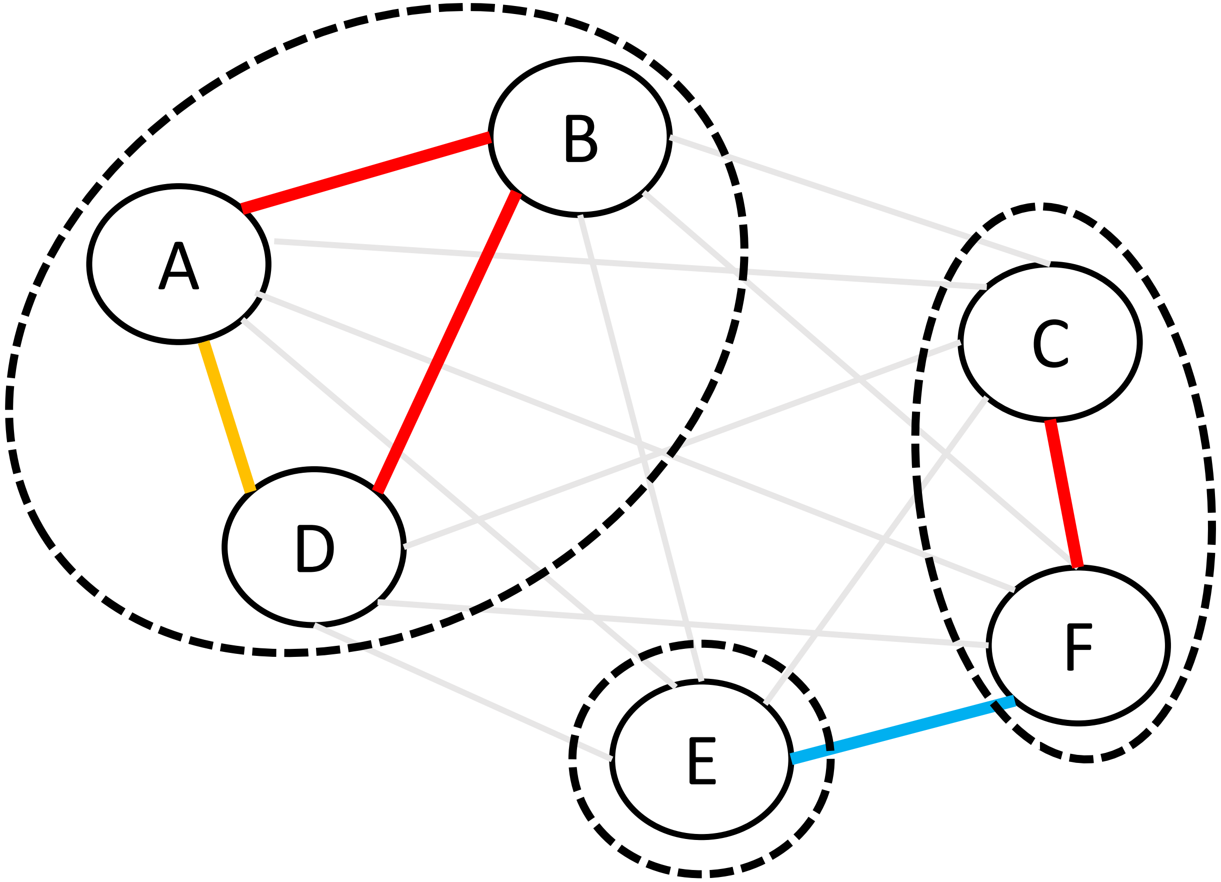
Case 2: all edges of E' are intra-category in $\mathcal{O}$



red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

Case 2: all edges of E' are intra-category in $\mathcal{O}$

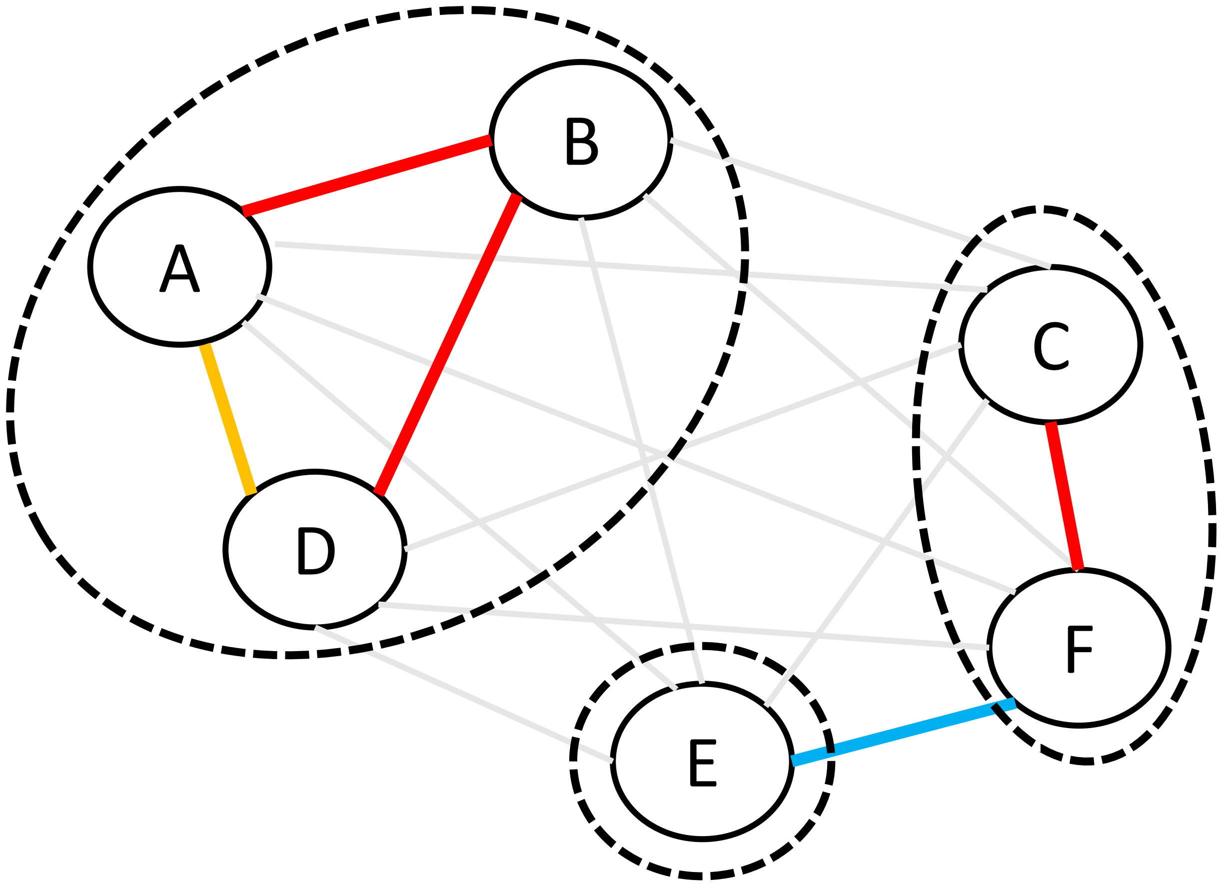


…Then $\mathcal{O}$ can't have **extra** intra-category edges that were inter-category in $\mathcal{G}$.

red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

Case 2: all edges of E' are intra-category in $\mathcal{O}$

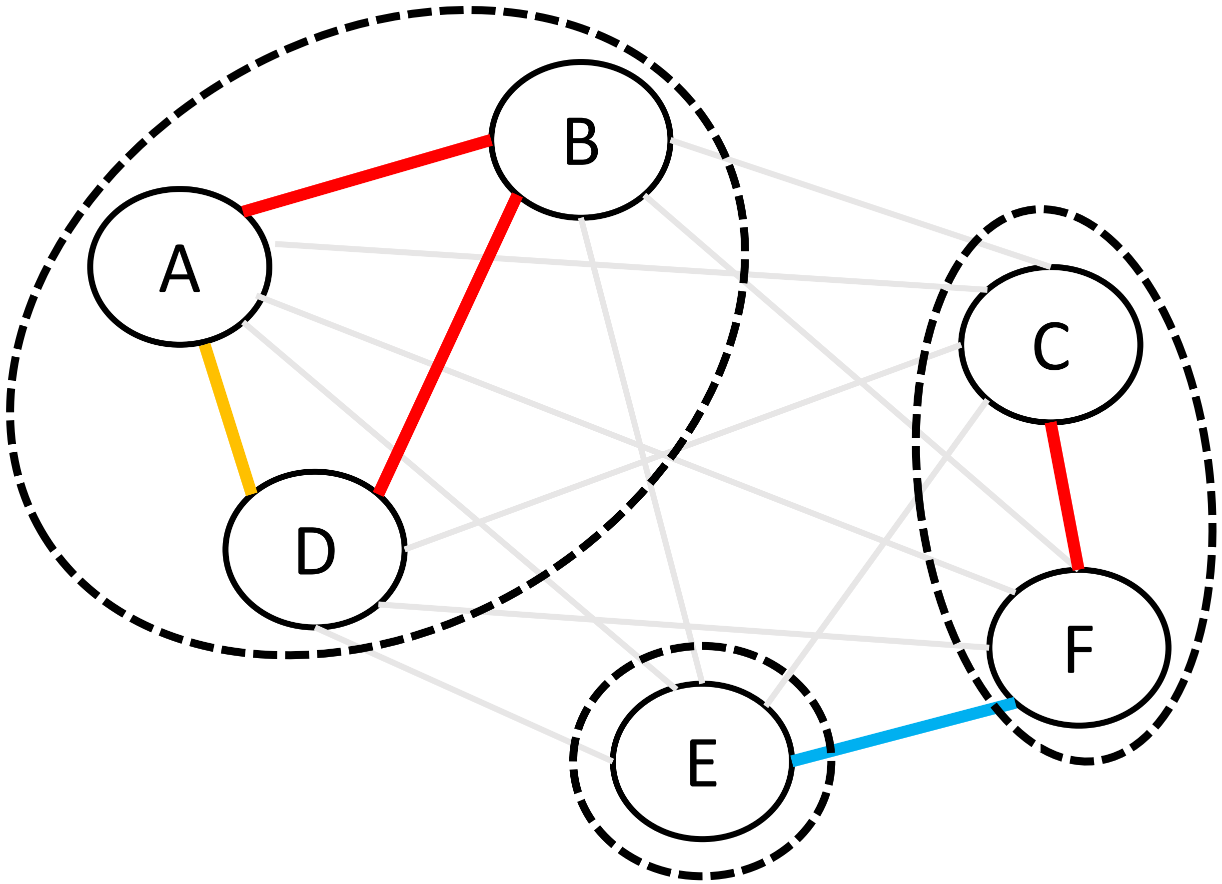


…Then $\mathcal{O}$ can't have **extra** intra-category edges that were inter-category in $\mathcal{G}$.
$\rightarrow$ all inter-category edges in $\mathcal{G}$ are inter-category in $\mathcal{O}$

red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

Case 2: all edges of E' are intra-category in $\mathcal{O}$



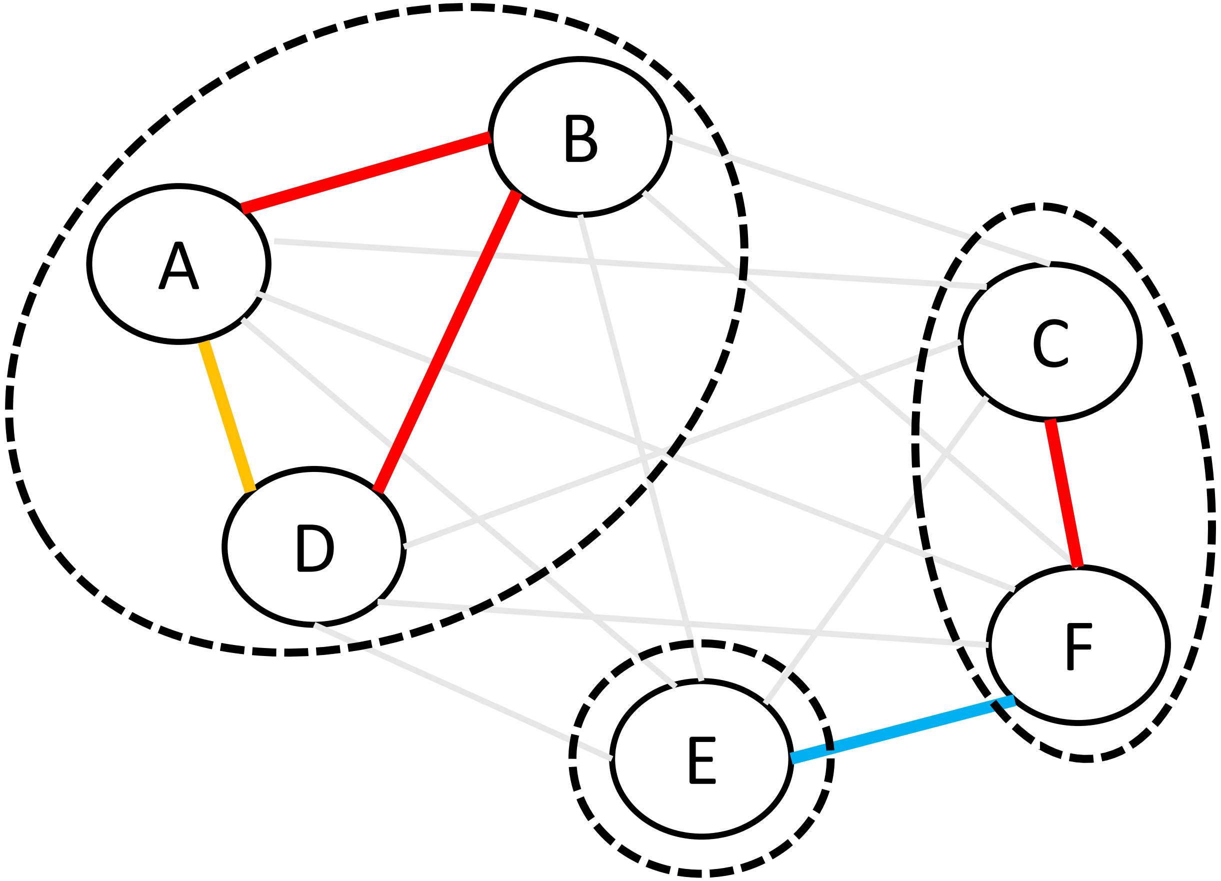
…Then $\mathcal{O}$ can't have **extra** intra-category edges that were inter-category in $\mathcal{G}$.
→ all inter-category edges in $\mathcal{G}$ are inter-category in $\mathcal{O}$*

*With a few extra steps, we can actually prove $\mathcal{G} = \mathcal{O}$

red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

Case 2: all edges of E' are intra-category in $\mathcal{O}$



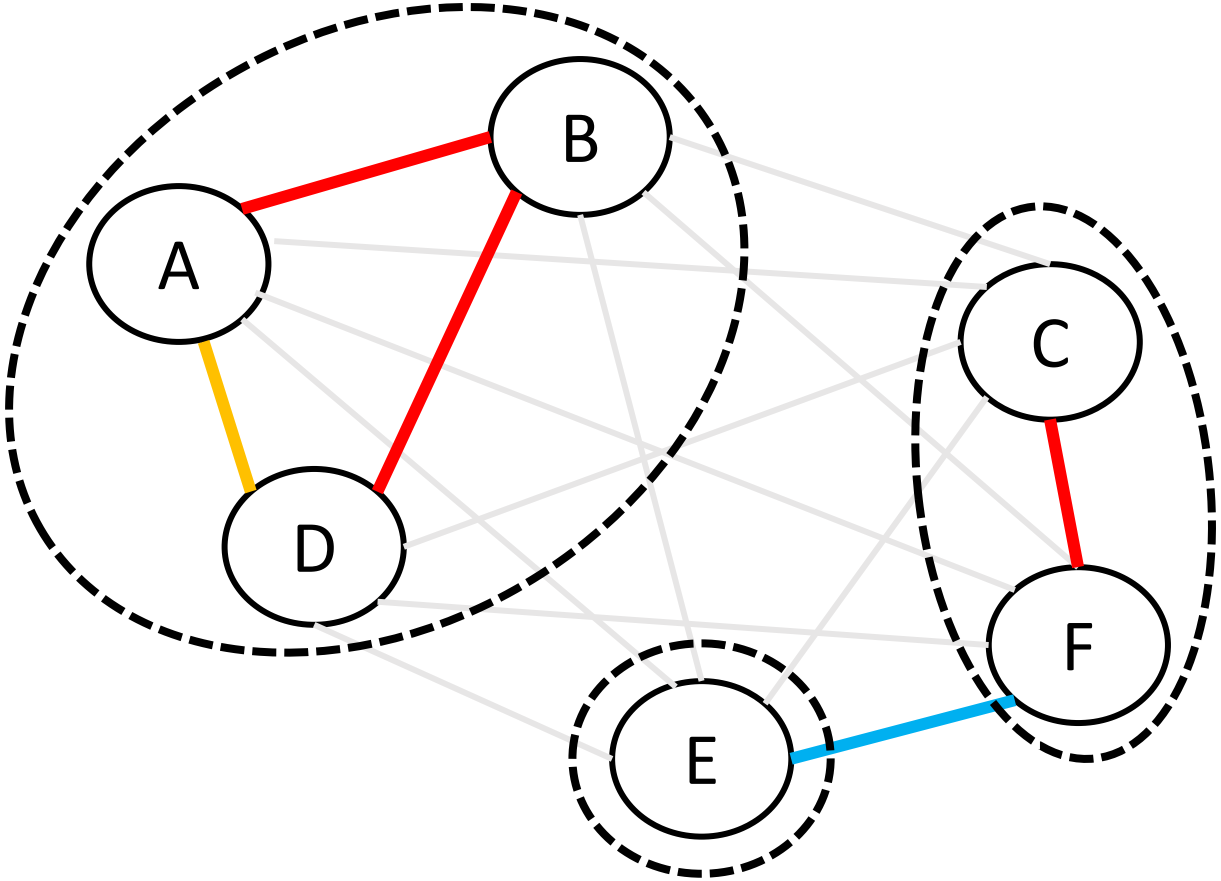
...Then $\mathcal{O}$ can't have **extra** intra-category edges that were inter-category in $\mathcal{G}$.
→ all inter-category edges in $\mathcal{G}$ are inter-category in $\mathcal{O}$*
→ Cost($\mathcal{G}$) ≤ Cost($\mathcal{O}$)

*With a few extra steps, we can actually prove $\mathcal{G} = \mathcal{O}$

red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

Summary

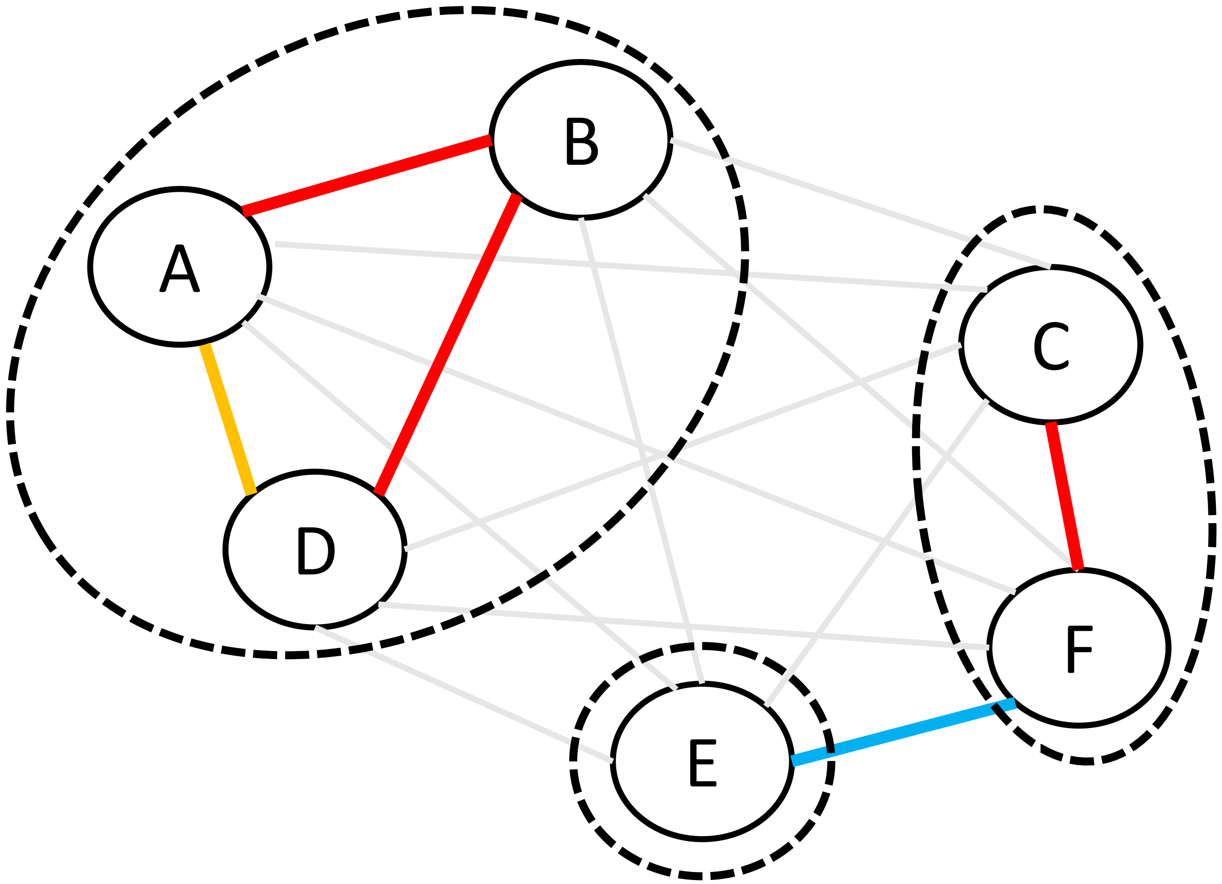


When some edges of E' are inter-category in $\mathcal{O}$:
$\text{Cost}(\mathcal{G}) \leq \text{Cost}(\mathcal{O})$

red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

Summary



When **some** edges of E' are inter-category in $\mathcal{O}$:

$$\text{Cost}(\mathcal{G}) \leq \text{Cost}(\mathcal{O})$$

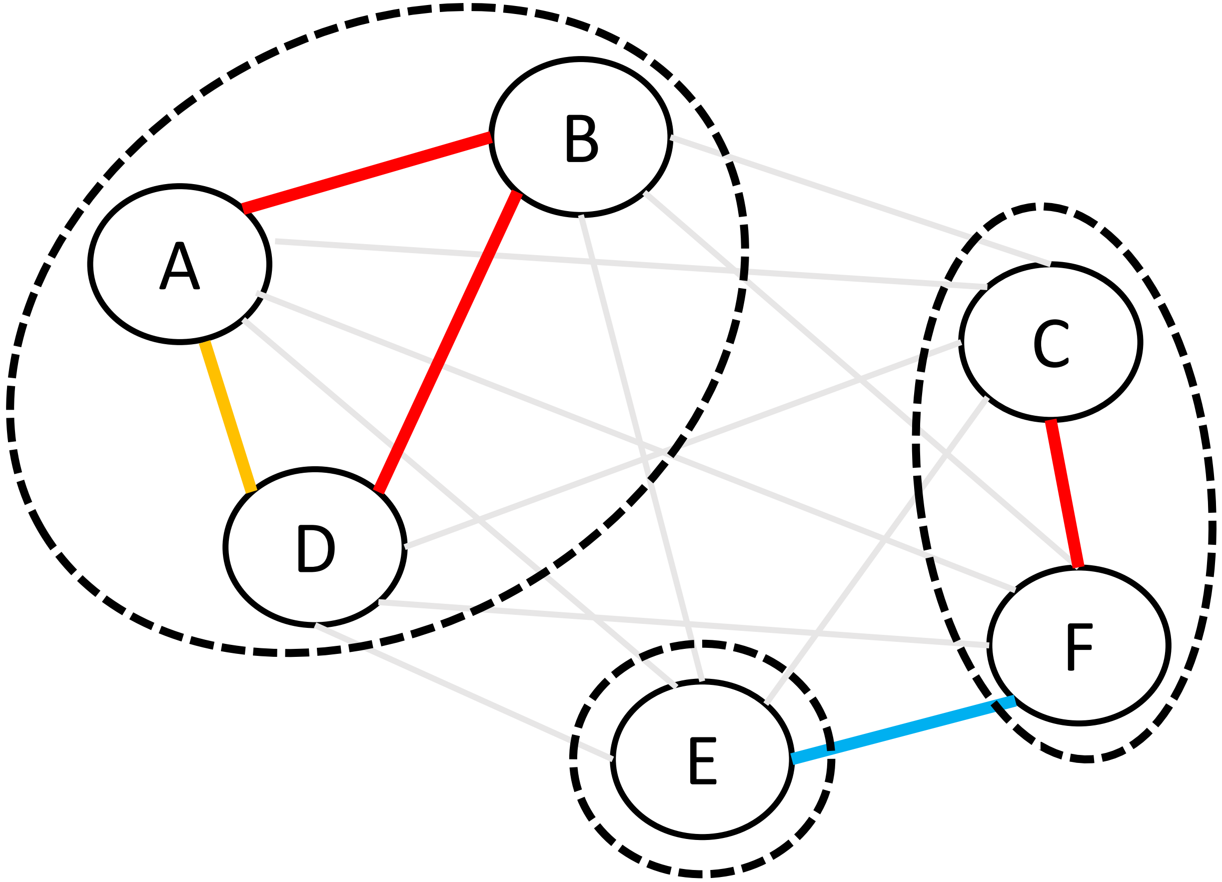
When **no** edges of E' are inter-category in $\mathcal{O}$:

$$\text{Cost}(\mathcal{G}) \leq \text{Cost}(\mathcal{O})$$

red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution

Summary



When **some** edges of E' are inter-category in $\mathcal{O}$:

$$\mathrm{Cost}(\mathcal{G}) \leq \mathrm{Cost}(\mathcal{O})$$

When **no** edges of E' are inter-category in $\mathcal{O}$:

$$\mathrm{Cost}(\mathcal{G}) \leq \mathrm{Cost}(\mathcal{O})$$

....QED, $\mathrm{Cost}(\mathcal{G}) \leq \mathrm{Cost}(\mathcal{O})$, which means $\mathcal{G}$ is optimal.

red edges = E' edges merged in Step 4
Orange = E' edges covered for free

The blue edge defines the cost of our greedy solution