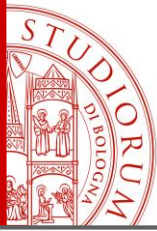


Cliques & Independent Sets

Alessandro Hill

rev. 1.0(AH) – 2024



The Maximum Clique Problem

Problema della cricca

Un sottoinsieme di nodi in V è chiamato "cricca" se tutti i nodi all'interno di questo sottoinsieme sono collegati tra loro.

Grafo non orientato

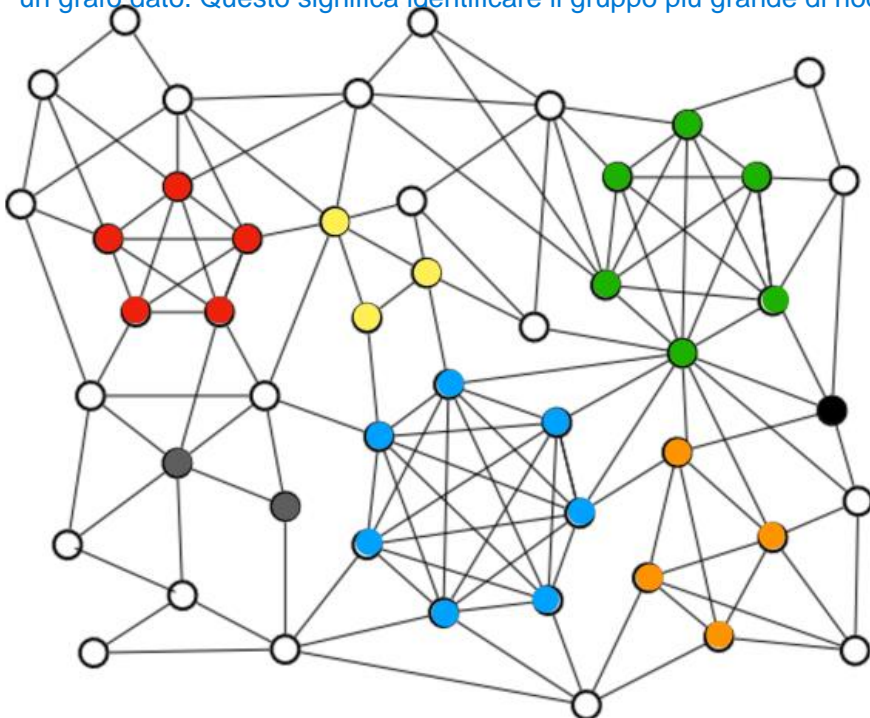
We are given an undirected base graph G with node set V and edge set E .

A subset of nodes in V is called a **clique** if all nodes are connected in G .

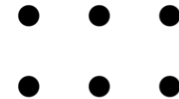
The **Maximum Clique Problem** asks for a clique of maximum cardinality in G .*

Trovare una cricca con il massimo numero di nodi (cardinalità massima) all'interno di un grafo dato. Questo significa identificare il gruppo più grande di nodi collegati tra loro.

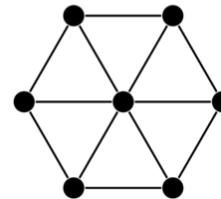
***NP-hard** (extremely difficult to optimize)



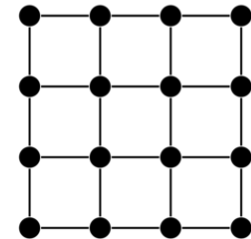
Claw



Empty graph



Wheel



Grid

Applications: Biology, social network analysis, telecommunications, computer science, etc.



The Maximum Clique Problem

IP Model*

Binary node variable for each node that could be part of a clique:

$$x_v = \begin{cases} 1 & \text{if node } v \text{ will be used in the clique,} \\ 0 & \text{otherwise.} \end{cases}$$

Maximize $\sum_{v \in V} x_v$

Massimizzare questa somma permette di trovare il sottoinsieme più grande di nodi che possono essere inclusi in una cricca.

non impostare il vincolo sulla diagonale della matrice di adiacenza dei nodi (perché auto legame con se stesso non può esistere)

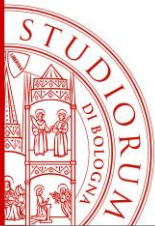
Idea: Two unconnected nodes cannot be both part of a clique.

$$x_u + x_v \leq 1$$

$$\forall \{u, v\} \in \bar{E}$$

sta a indicare che il vincolo si applica a tutte le coppie di nodi u e v che non sono connessi nel grafo.

Questo vincolo garantisce che solo nodi connessi tra loro possano appartenere simultaneamente alla cricca, mantenendo la definizione di "cricca" (ossia, un sottoinsieme di nodi completamente connesso).



Vado a prendere quei nodi con maggior vicini e collegati con quelli già scelti (perché vado a prendere quelli che hanno maggior vicini insieme a quelli già scelti, cioè già presenti in Q)

The Maximum Clique Problem

Max-Degree Heuristic (Greedy)

Grafo non orientato

Input: Undirected graph $G=(V,E)$.

Degree = Grado

1. Initialize clique $Q = \{\}$.

2. Pick a node v in V of highest degree. Add v to Q .

Scegli un nodo v in V che abbia il grado massimo (ossia che sia connesso al maggior numero di altri nodi nel grafo). Aggiungilo a Q .

3. Pick the node v in $V \setminus Q$ that

- is connected to all nodes in Q and
- has the highest number of neighbors in common with nodes in Q .

Add v to Q .

4. If no node v was found, return Q , else go to 3.

L'euristica tenta di costruire la cricca massima aggiungendo, a ogni passo, il nodo più connesso che è compatibile con la struttura della cricca formata fino a quel momento. Anche se questa tecnica non garantisce sempre la soluzione ottimale (poiché è un metodo greedy), è spesso efficace e veloce per trovare una soluzione approssimata, specialmente per grafi di grandi dimensioni.

A questo punto, cerchiamo di aggiungere nuovi nodi a Q seguendo i criteri specificati: Scegli un nodo v in $V \setminus Q$ (cioè un nodo che non è ancora stato aggiunto a Q) che:

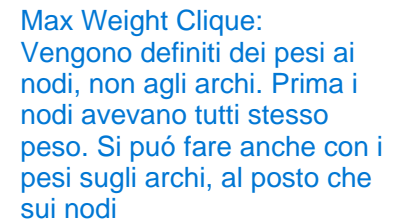
- 1) Sia connesso a tutti i nodi già presenti in Q .
- 2) Abbia il massimo numero di vicini in comune con i nodi in Q .

Aggiungi questo nodo v a Q .



Example:

Per escludere una Clique (cioè selezionati quei nodi insieme) dalle disponibili, applico un vincolo dove faccio la somma dei nodi selezionati minore o uguale a cardinalità dei selezionati - 1



- <http://listofrandomnames.com>

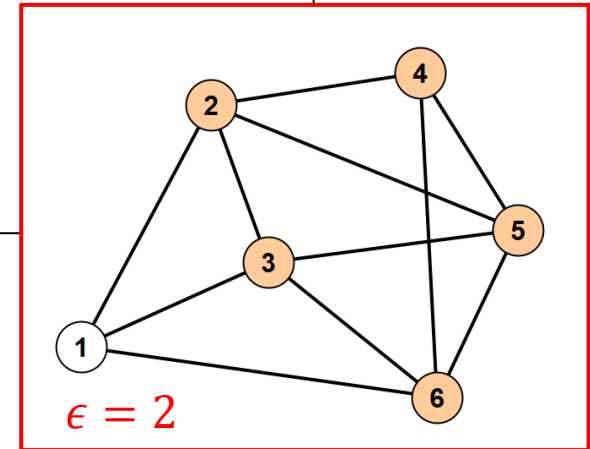
The Maximum Quasi-Clique Problem

An ϵ -quasi-clique, is a clique that is missing ϵ edges.

Binary “exception variable” for unconnected nodes $\{u, v\} \in \bar{E}$:

$$z_{u,v} = \begin{cases} 1 & \text{if unconnected nodes } u \text{ and } v \text{ are} \\ & \text{part of the quasi-clique,} \\ 0 & \text{otherwise.} \end{cases}$$

se due nodi non connessi (u e v) sono comunque considerati parte del quasi-clique.



Maximize $\sum_{v \in V} x_v$

x_v è una variabile che indica se il nodo v appartiene al quasi-clique

Idea: Two unconnected nodes cannot be both part of a clique unless they form an exception.

sta a indicare che il vincolo si applica a tutte le coppie di nodi u e v che non sono connessi nel grafo.

$$x_u + x_v \leq 1 + z_{u,v}$$

$$\forall \{u, v\} \in \bar{E}$$

Due nodi non connessi (u e v) possono essere inclusi nella stessa quasi-clique solo se accettiamo questa come un'eccezione.

Limit the number of exceptions to ϵ :

$$\sum_{\{u,v\} \in \bar{E}} z_{u,v} \leq \epsilon$$

Per garantire che il quasi-clique rispetti la definizione di epsilon-quasi-clique, il numero totale di eccezioni non deve superare epsilon



The Maximum Independent Set Problem

Problema del massimo insieme indipendente

We are given an undirected base graph G with node set V and edge set E .

A subset of nodes in V is called an **independent set** if all nodes are disconnected.

Tutti i nodi del sottoinsieme sono non connessi tra loro

The **Maximum Independent Set Problem** asks for an independent set of maximum cardinality in G . *

Se aggiungi un nodo a un Maximal Clique o a un Maximal Independent Set, non sarà più un clique o un insieme indipendente. La massimalità riguarda l'impossibilità di aggiungere nodi senza violare le regole

Questa relazione consente di risolvere il problema sfruttando algoritmi per il problema dei clique.

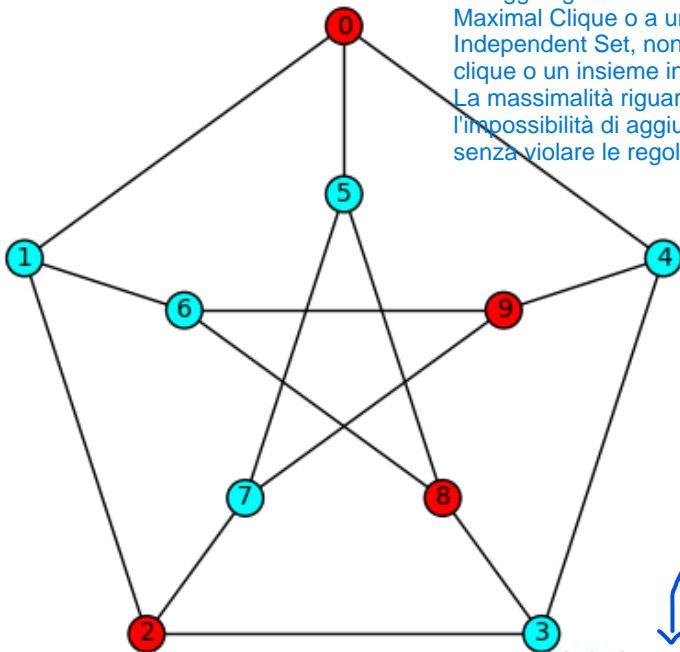
*NP-hard (extremely difficult to optimize)

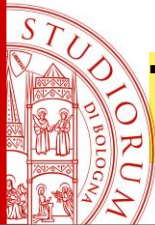
Observation: An independent set of G is a clique in its complement graph \bar{G} .

Consequence: A maximum clique in \bar{G} is a maximum independent set in G .

Definition: A clique (independent set) is called **maximal** if it cannot be extended by any node.

Un insieme indipendente è detto massimale se non è possibile aggiungere altri nodi senza violare la proprietà di indipendenza. Un clique è massimale se non è possibile aggiungere altri nodi senza perdere la proprietà di connessione.





The Maximum Independent Set Problem

IP Model*

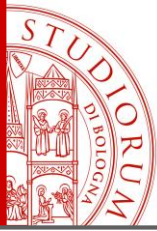
Binary node variables for nodes that could be part of an independent set:

$$x_v = \begin{cases} 1 & \text{if node } v \text{ will be used in the independent set,} \\ 0 & \text{otherwise.} \end{cases}$$

Maximize $\sum_{v \in V} x_v$

Idea: Two connected nodes cannot be both part of an independent set.

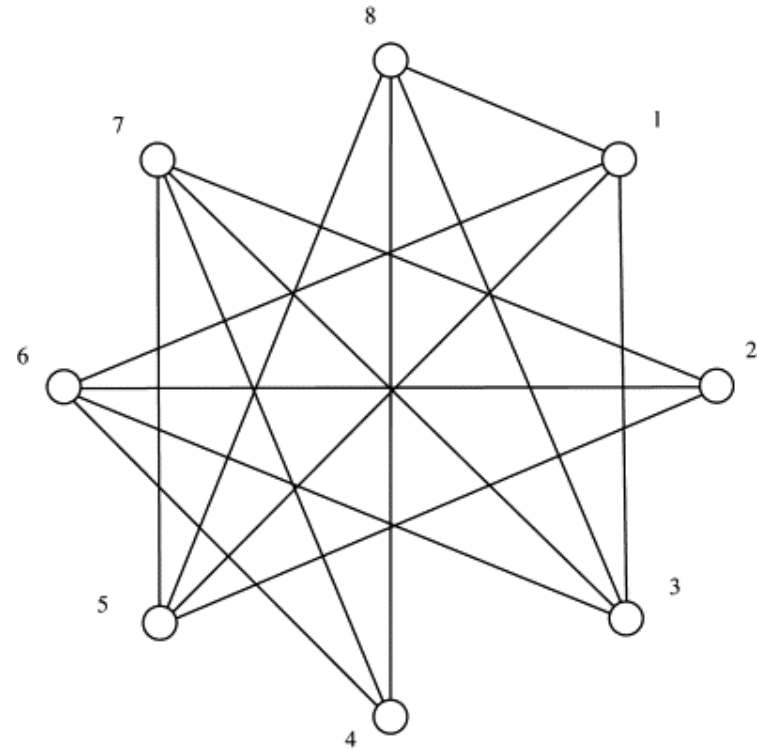
$$x_u + x_v \leq 1 \quad \forall \{u, v\} \in E$$



Cliques & Independent Sets

1) Exercises:

1. Consider the graph with 8 nodes.
2. Find a maximum clique using IP.
Verify your results visually.
3. Find a maximum clique using the max-degree heuristic.
Verify your results visually.
4. Find all maximum cliques using IP.
Verify your results visually.
5. Find a maximum quasi-clique using IP.
Use $\epsilon \in \{1,2,3,4\}$.
Verify your results visually.
6. Find a maximum independent set using IP.
Verify your results visually.





Cliques & Independent Sets

2) Exercises:

- 1. Create a random graph with 20 nodes and 80 edges in yEd.**
Use a convenient layout for your visualization.
Format as necessary.
You can export the edge list using the TGF format.
- 2. Find a maximum clique using IP.**
Verify your results visually.
- 3. Find a maximum quasi-clique using IP.**
Use $\epsilon \in \{1,5,10\}$.
Verify your results visually.
- 4. Find a maximum independent set using IP.**
Verify your results visually.
- 5. Can you find two disjoint cliques such that the overall number of selected clique nodes is maximized?**
Use IP.
- 6. For what larger graph sizes/densities can you still answer the questions above?**

230	9	80
231	9	81
232	10	21
233	10	38
234	10	60
235	10	82
236	10	83
237	10	104
238	10	140
239	10	68
240	10	80
241	10	81
242	11	12
243	11	13
244	11	14
245	11	15
246	11	54
247	11	84
248	11	85
249	12	13
250	12	14
251	12	15
252	12	53
253	12	54
254	12	55
255	12	84
256	12	85
257	12	41
258	12	144
259	12	143
260	13	14
261	13	15
262	13	53
263	13	54
264	13	55
265	13	84
266	13	85
267	13	144
268	13	143
269	14	15
270	14	53
271	14	55
272	14	84
273	14	85
274	14	143
275	14	144
276	15	53
277	15	54
278	15	52
279	15	55



Cliques & Independent Sets

3) Exercises:

1. Use the class social network data (Virtuale).
2. Find a maximum clique using IP that includes yourself.
Use arbitrary node number if you cannot remember yours.
Verify your results visually.
3. Find a maximum independent set using IP that includes you.
Use arbitrary node number if you cannot remember yours.
Verify your results visually.

