Cliques & Independent Sets

Alessandro Hill

rev. 1.0(AH) – 2024

The Maximum Clique Problem

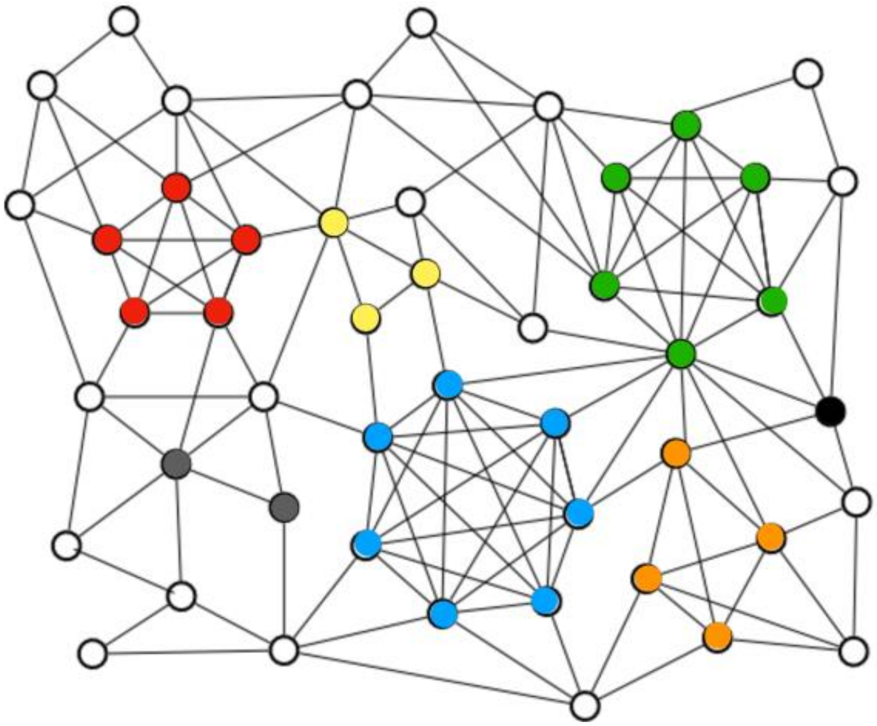
Problema della cricca

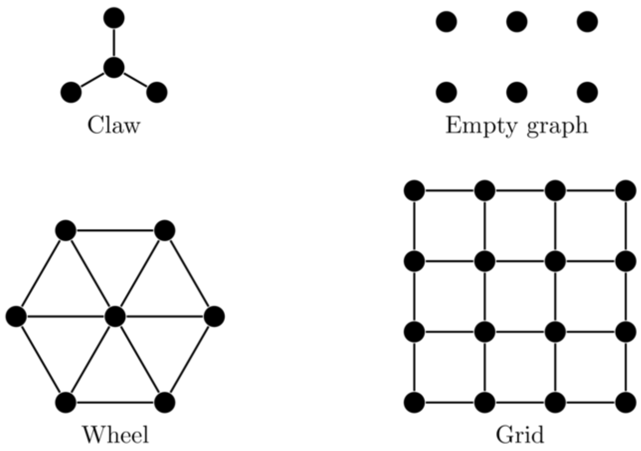
We are given an undirected base graph **G** with node set **V** and edge set **E**.

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**.*

*NP-hard (extremely difficult to optimize)





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:

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

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

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

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

The Maximum Clique Problem

Max-Degree Heuristic (Greedy)

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

1. Initialize clique Q = {}.

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

3. Pick the node v in V\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.

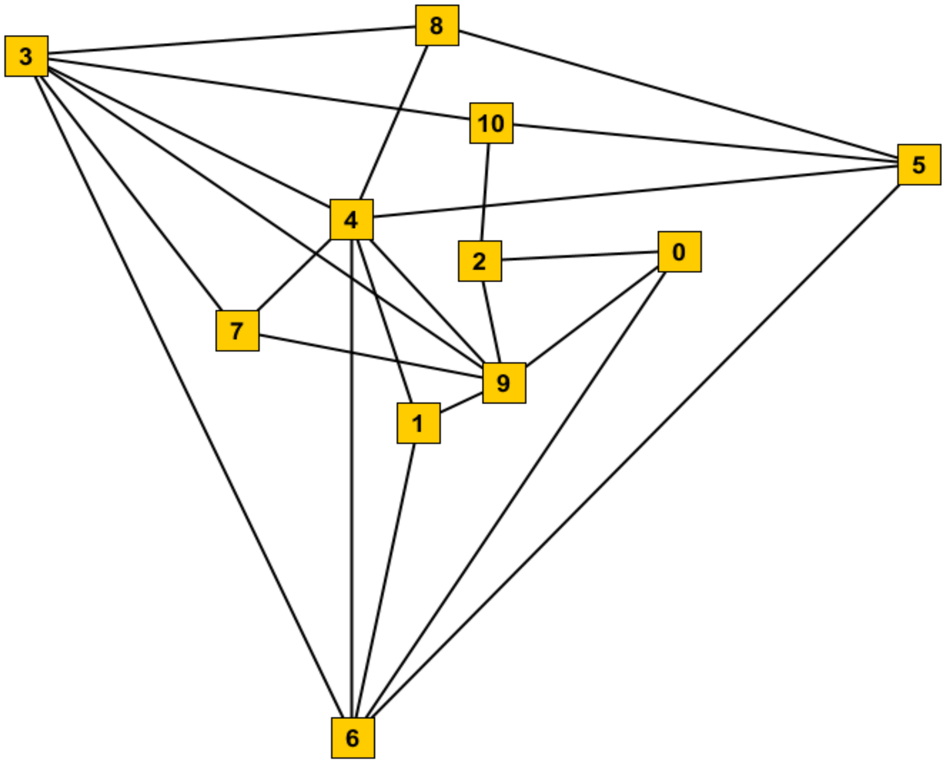
The Maximum Clique Problem

What is the largest clique the social network given below?

Example:

ID	Name	Friends With
0	Toi	2,9,6
1	Brain	4,6,9
2	Annamaria	0,9,10
3	Nina	4,6,7,8,9,10
4	Walton	1,3,5,6,7,8,9
5	Virgilio	4,6,8,10
6	Teena	0,1,3,4,5
7	Darrin	3,4,9
8	Alessandra	3,4,5
9	Harry	0,1,2,3,4,7
10	Simona	2,3,5

http://listofrandomnames.com



a) Find **a** maximum clique using **IP**.

b) Find **ALL** maximum cliques using **IP**.

c) Find a clique using the maximum-degree heuristic.

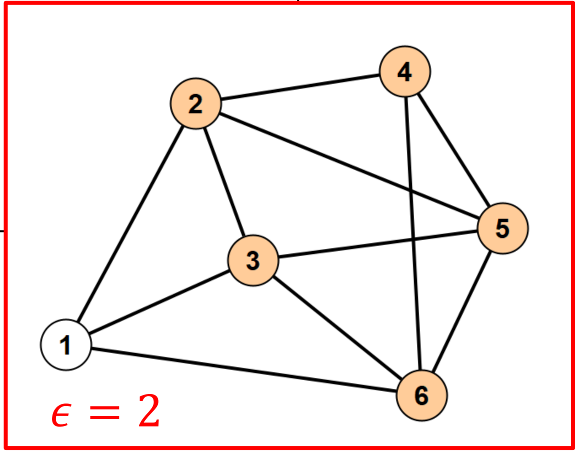
d) Find a maximum weight clique using node weight ‘**nodeIndex**’ (**IP**).

The Maximum Quasi-Clique Problem

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

Binary "exception variable" for unconnected nodes $\{u, v\} \in \overline{E}$:

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



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

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

$$x_u + x_v \leq 1 + z_{u,v} \qquad \forall \{u, v\} \in \overline{E}$$

Limit the number of exceptions to ϵ:

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

The Maximum Independent Set Problem

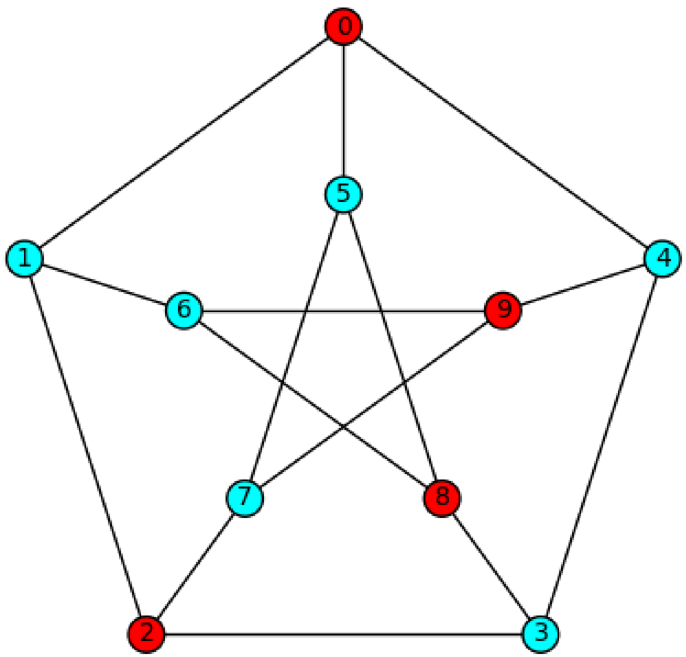
Problema del massimo insieme indipendente

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.

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

*NP-hard (extremely difficult to optimize)



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

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

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

The Maximum Independent Set Problem

IP Model*

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

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

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

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

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

Cliques & Independent Sets

1) Exercises:

1. Consider the graph with 8 nodes.

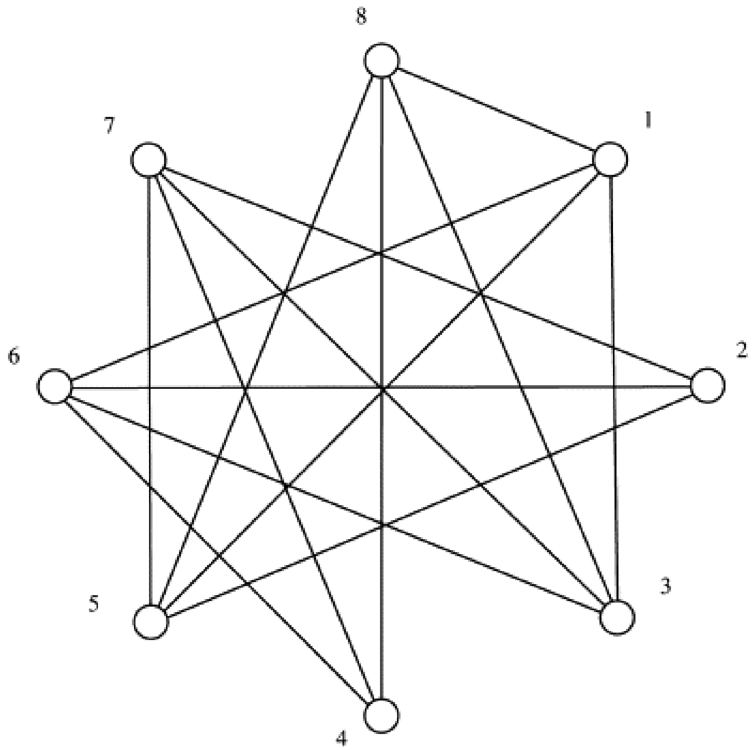
2. Find a maximum clique using IP.
Verify your results visually.

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.

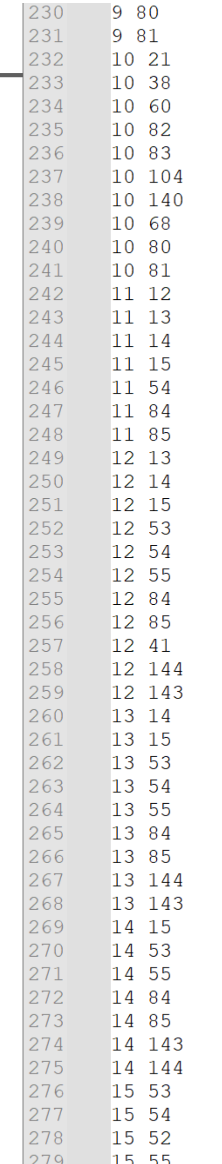
2. Find a maximum clique using IP.
Verify your results visually.

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 sum of is maximized?
Use IP.

6. For what larger graph sizes can you still answer the questions above?



```
9 80
9 81
10 21
10 38
10 60
10 82
10 83
10 104
10 140
10 68
10 80
10 81
11 12
11 13
11 14
11 15
11 54
11 84
11 85
12 13
12 14
12 15
12 53
12 54
12 55
12 84
12 85
12 41
12 144
12 143
13 14
13 15
13 53
13 54
13 55
13 84
13 85
13 144
13 143
14 15
14 53
14 55
14 84
14 85
14 143
14 144
15 53
15 54
15 52
15 55
```

Cliques & Independent Sets

3) Exercises:

1. Use the class social network data (Virtuale).

2. Find a maximum clique using IP that includes you.
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.