# Graph Theory and Complex Networks: An Introduction

Maarten van Steen

VU Amsterdam, Dept. Computer Science
Room R4.20, steen@cs.vu.nl

## Chapter 02: Foundations

Version: April 8, 2014

vrije Universiteit amsterdam

# Contents

| Chapter | Description |
|---|---|
| 01: Introduction | History, background |
| **02: Foundations** | Basic terminology and properties of graphs |
| 03: Extensions | Directed & weighted graphs, colorings |
| 04: Network traversal | Walking through graphs (cf. traveling) |
| 05: Trees | Graphs without cycles; routing algorithms |
| 06: Network analysis | Basic metrics for analyzing large graphs |
| 07: Random networks | Introduction modeling real-world networks |
| 08: Computer networks | The Internet & WWW seen as a huge graph |
| 09: Social networks | Communities seen as graphs |

# Graph: definition

**Definition**

A graph $G$ is a tuple $(V, E)$ of vertices $V$ and a collection of edges $E$. Each edge $e \in E$ is said to connect two vertices $u, v \in V$, and is denoted as $e = \langle u, v \rangle$.
Notations: $V(G)$, $E(G)$.

**Definition**

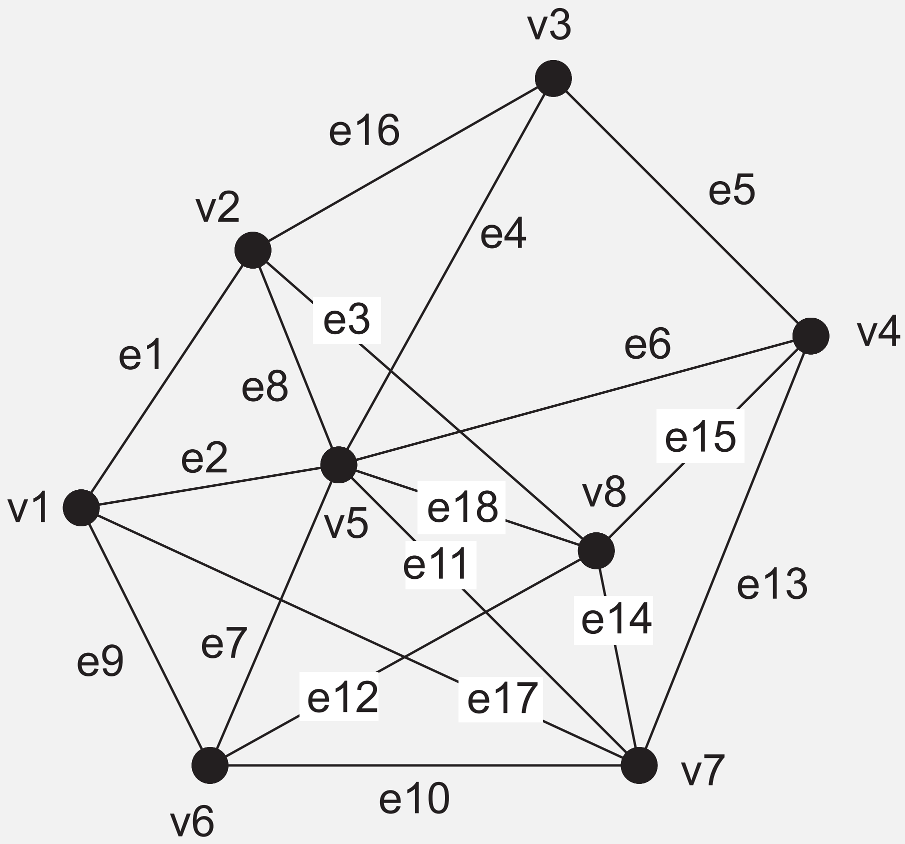
The complement $\overline{G}$ of a graph $G$, has the same vertex set as $G$, but $e \in E(\overline{G})$ *if and only if* $e \notin E(G)$.

**Definition**

For any graph $G$ and vertex $v \in V(G)$, the neighbor set $N(v)$ of $v$ is the set of vertices (other than $v$) adjacent to $v$:

$$N(v) = \{w \in V(G) \mid v \neq w, \langle v, w \rangle \in E(G)\}$$

# Graph: Example




$$V(G) = \{v_1, \ldots, v_8\}$$
$$E(G) = \{e_1, \ldots, e_{18}\}$$

$$\begin{array}{ll} e_1 = \langle v_1, v_2 \rangle & e_{10} = \langle v_6, v_7 \rangle \\ e_2 = \langle v_1, v_5 \rangle & e_{11} = \langle v_5, v_7 \rangle \\ e_3 = \langle v_2, v_8 \rangle & e_{12} = \langle v_6, v_8 \rangle \\ e_4 = \langle v_3, v_5 \rangle & e_{13} = \langle v_4, v_7 \rangle \\ e_5 = \langle v_3, v_4 \rangle & e_{14} = \langle v_7, v_8 \rangle \\ e_6 = \langle v_4, v_5 \rangle & e_{15} = \langle v_4, v_8 \rangle \\ e_7 = \langle v_5, v_6 \rangle & e_{16} = \langle v_2, v_3 \rangle \\ e_8 = \langle v_2, v_5 \rangle & e_{17} = \langle v_1, v_7 \rangle \\ e_9 = \langle v_1, v_6 \rangle & e_{18} = \langle v_5, v_8 \rangle \end{array}$$

**Question**

What is the neighborset of $v_6$?

# Vertex degree

**Definition**

The number of edges incident with a vertex $v$ is called the degree of $v$, denoted as $\delta(v)$. Loops, i.e., edges joining a vertex with itself, are counted twice.

**Theorem**

*For all graphs G,* $\sum_{v \in V(G)} \delta(v)$ *is* $2 \cdot |E(G)|$.

**Proof**

When we count the edges of a graph $G$ by enumerating the edges incident with each vertex of $G$, we are counting each edge exactly twice.

# Degree sequence

## Definition

An (ordered) degree sequence is an (ordered) list of the degrees of the vertices of a graph. A degree sequence is graphic if there is a (simple) graph with that sequence.

## Theorem (Havel-Hakimi)

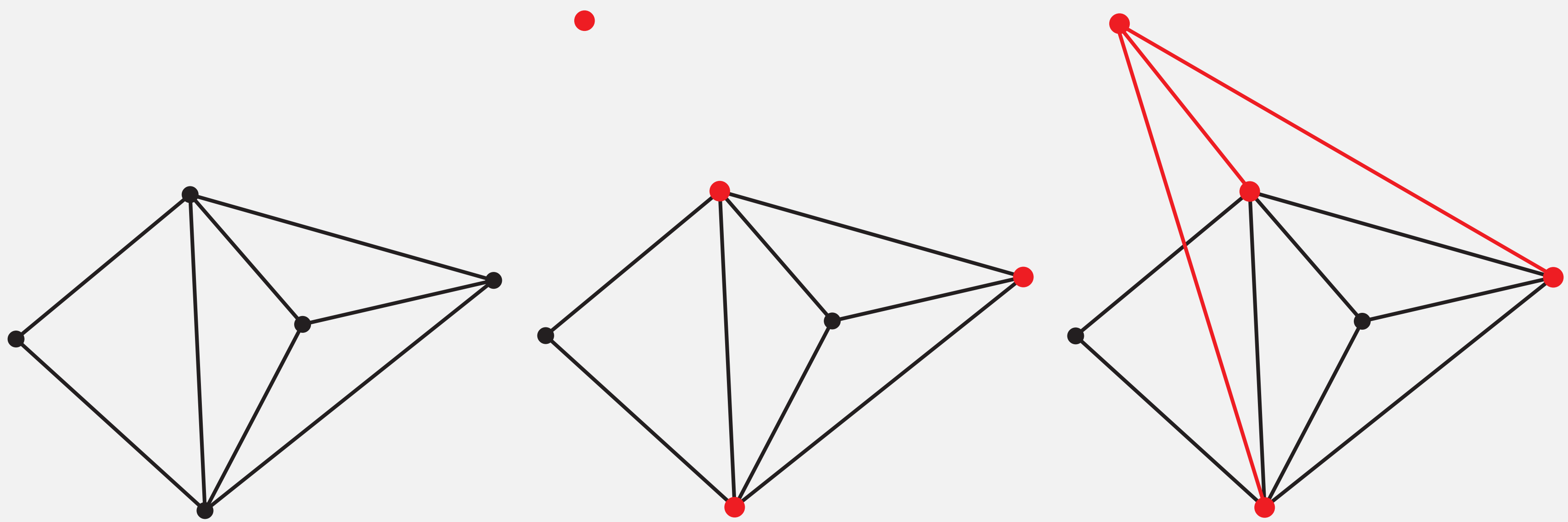
*An ordered degree sequence* $\mathbf{s} = [k, d_1, d_2, \ldots, d_{n-1}]$ *is graphic, if and only if* $\mathbf{s}^* = [d_1 - 1, d_2 - 1, \ldots, d_k - 1, d_{k+1}, \ldots, d_{n-1}]$ *is also graphic. (We assume* $k \geq d_i \geq d_{i+1}$*.)*

## Note

Length $\mathbf{s} = n$, but length $\mathbf{s}^* = n - 1$.

# $\mathbf{s}^* \Rightarrow \mathbf{s}$: Example

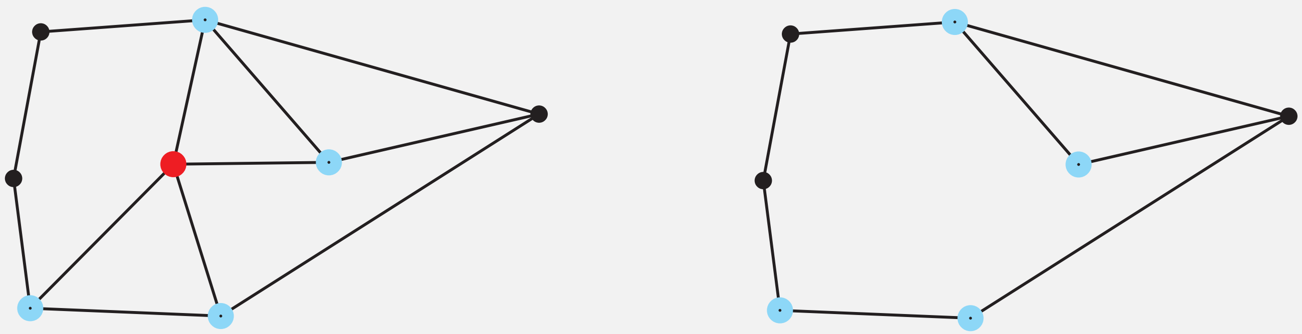
Take $k = 3$ and consider graph with sequence $[4,4,3,3,2]$. Create graph with sequence $[3,5,5,4,3,2] \equiv [5,5,4,3,3,2]$:



1. Starting condition
2. Add a vertex $v$ with degree $\delta(v) = k$
3. Connect $v$ to $k$ vertices with highest degrees.

# $\mathbf{s} \Rightarrow \mathbf{s}^*$: Example

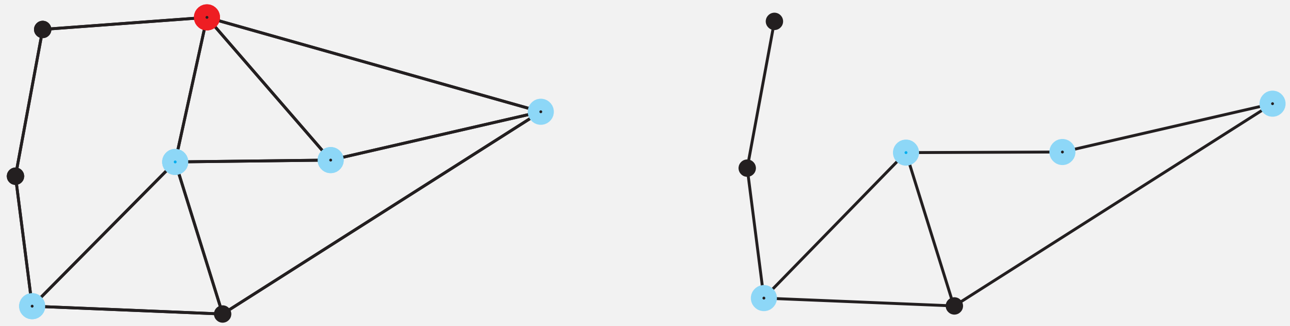
Consider the following graph with sequence $[4,4,3,3,3,3,2,2]$. Let $\delta(u) = 4$ (in red) and consider $V = \{v_1, v_2, v_3, v_4\}$ as next highest degrees (in blue), and $W = \{w_1, w_2, w_3\}$ the rest (in black).



1. Starting condition
2. Remove $u$. Because $u$ is connected only to vertices from $V$, we know that $s^* = [3,2,2,2,3,2,2] = [3,3,2,2,2,2,2]$

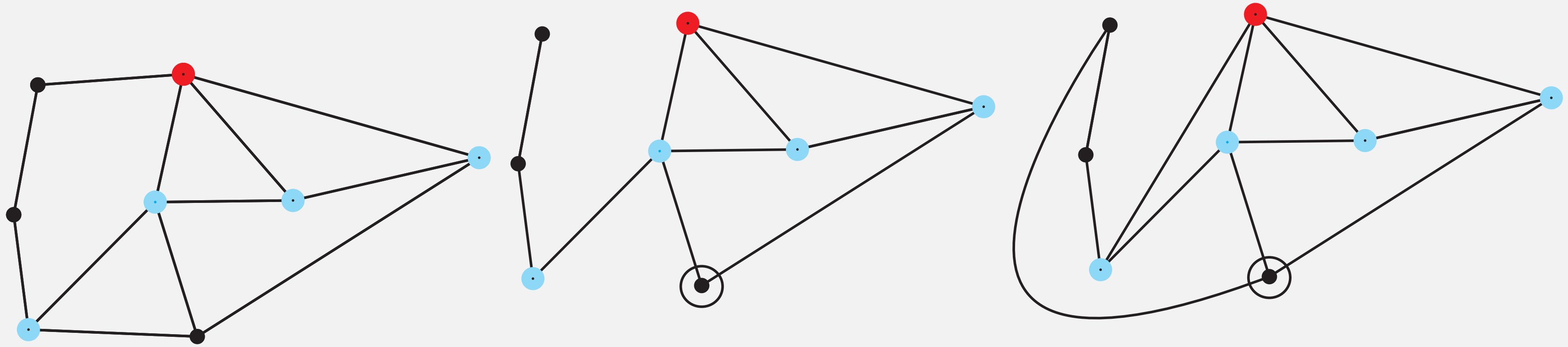
# $\mathbf{s} \Rightarrow \mathbf{s}^*$: Example

Consider the following graph with sequence $[4,4,3,3,3,3,2,2]$. Let $\delta(u) = 4$ (in red) and consider $V = \{v_1, v_2, v_3, v_4\}$ as next highest degrees (in blue), and $W = \{w_1, w_2, w_3\}$ the rest (in black).



1. Starting condition
2. Remove $u$. Because $u$ is **not** connected only to vertices from $V$, we have a problem: $s^* = [3,3,3,2,2,2,1]$.

# $\mathbf{s} \Rightarrow \mathbf{s}^*$: Example



1. Problem: $u$ is linked to a $w$ but not to a $v_j$, with $\delta(w) < \delta(v_j)$. But because $\delta(w) < \delta(v_j)$, there exists $x$ adjacent to $v_j$ but not to $w$.
2. Remove $\langle u, w \rangle$ and $\langle v_j, x \rangle$.
3. Add $\langle x, w \rangle$ and $\langle u, v_j \rangle$

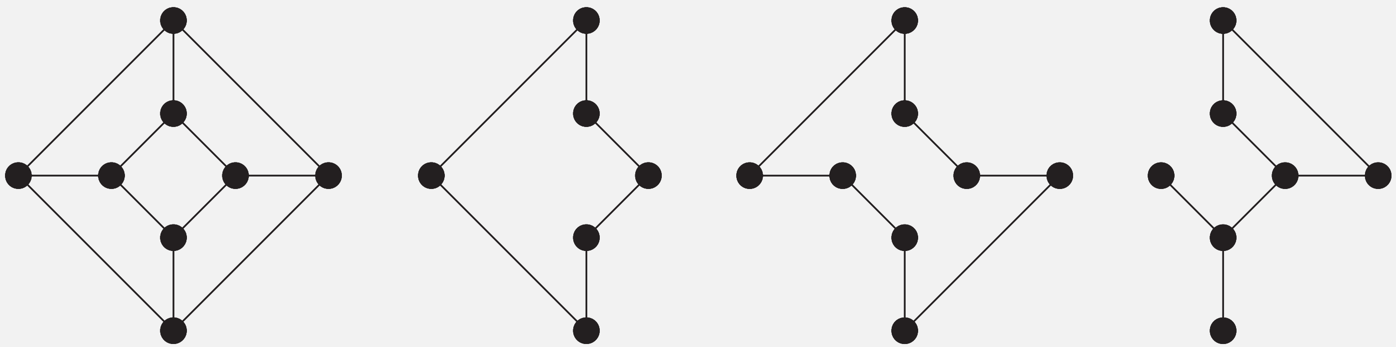
**Question**

What should we do if $u$ was linked to a $w$ with $\delta(w) = \delta(v_j)$?

# Subgraphs

## Definition

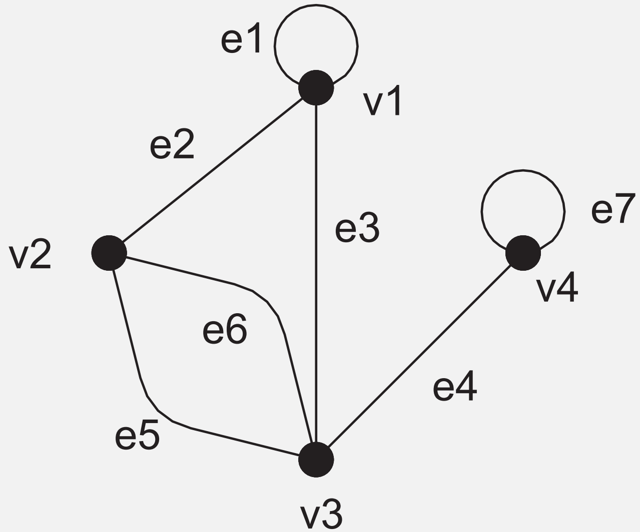
$H$ is a subgraph of $G$ if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$ such that for all $e \in E(H)$ with $e = \langle u, v \rangle : u, v \in V(H)$.



## Definition

The subgraph induced by $V^* \subseteq V(G)$ has vertex set $V^*$ and edge set $\{\langle v, w \rangle \in E(G) | v, w \in V^*\}$. Denoted as $H = G[V^*]$. The subgraph induced by $E^* \subseteq E(G)$ has vertex set $V(G)$ and edge set $E^*$. Denoted as $H = G[E^*]$.

# Adjacency matrix

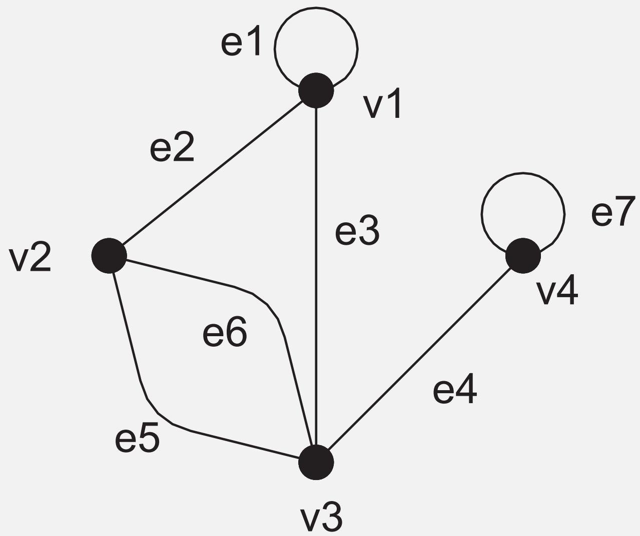



| | $v_1$ | $v_2$ | $v_3$ | $v_4$ |
|---|---|---|---|---|
| $v_1$ | 2 | 1 | 1 | 0 |
| $v_2$ | 1 | 0 | 2 | 0 |
| $v_3$ | 1 | 2 | 0 | 1 |
| $v_4$ | 0 | 0 | 1 | 2 |

## Observations

- Adjacency matrix is *symmetric*: $\mathbf{A}[i,j] = \mathbf{A}[j,i]$.
- $G$ is simple $\Leftrightarrow$ $\mathbf{A}[i,j] \leq 1$ and $\mathbf{A}[i,i] = 0$.
- $\forall v_i$: $\sum_{j=1}^{n} \mathbf{A}[i,j] = \delta(v_i)$.

# Incidence matrix




| | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
|---|---|---|---|---|---|---|---|
| $v_1$ | 2 | 1 | 1 | 0 | 0 | 0 | 0 |
| $v_2$ | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| $v_3$ | 0 | 0 | 1 | 1 | 1 | 1 | 0 |
| $v_4$ | 0 | 0 | 0 | 1 | 0 | 0 | 2 |

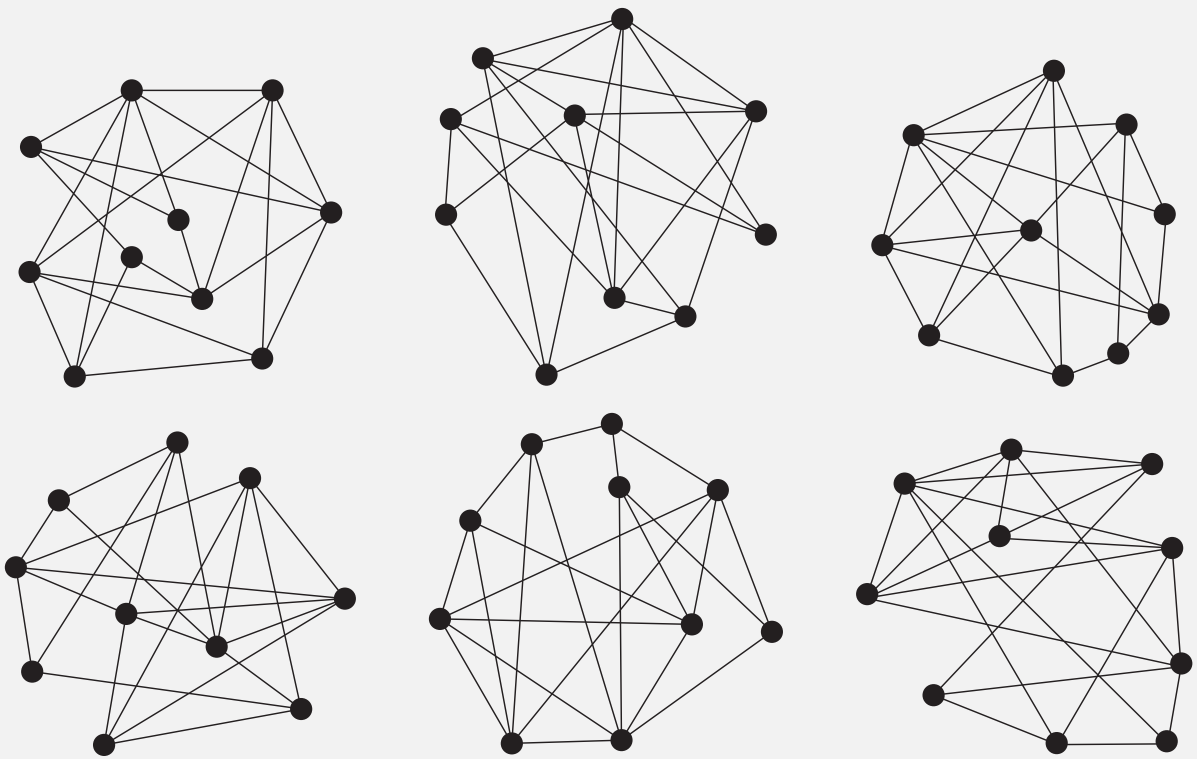
## Observations

- $G$ is simple only if $\mathbf{M}[i,j] \leq 1$
- $\forall v_i$: $\sum_{j=1}^{m} \mathbf{M}[i,j] = \delta(v_i)$.
- $\forall e_j$: $\sum_{i=1}^{n} \mathbf{M}[i,j] = 2$.

# Graph isomorphism

**Definition**

$G_1$ and $G_2$ are isomorphic if there exists a one-to-one mapping $\phi : V_1 \to V_2$ such that for each edge $e_1 \in E_1$ with $e_1 = \langle v, u \rangle$ there is a unique edge $e_2 \in E_2$ with $e_2 = \langle \phi(v), \phi(u) \rangle$.



# Connectivity: definitions

**Definition**

A $(\mathbf{v_0}, \mathbf{v_k})$-walk is a sequence $[v_0, e_1, v_1, e_2, \ldots, v_{k-1}, e_k, v_k]$ with $e_i = \langle v_{i-1}, v_i \rangle$. A trail is a walk with distinct edges; a path is a trail with distinct vertices. A cycle is a trail with distinct vertices except $v_0 = v_k$.

**Definition**

Vertices $u \neq v$ in $G$ are connected if there is a $(u, v) - path$ in $G$. $G$ is connected if all pairs of distinct vertices are connected.

**Definition**

$H \subseteq G$ is a component of $G$ if $H$ is connected and not contained in a connected subgraph of $G$ with more vertices or edges. The number of components of $G$ is $\omega(G)$.

# Connectivity and robustness

**Important**

Connectivity indicates whether all nodes in a network can be reached from any other node.

**Example**

Communication networks, like the Internet, require to be connected, and have been designed to stay connected, even when under attack.

**Definition**

For a graph $G$ let $V^* \subset V(G)$ and $E^* \subset E(G)$. If $\omega(G - V^*) > \omega(G)$ then $V^*$ is called a vertex cut. If $\omega(G - E^*) > \omega(G)$ then $E^*$ is called an edge cut.

# Minimal cuts

## Note

For reasons of robustness, we're interested in finding the minimal number of vertices or edges to remove before a graph falls apart.

## Notations

- $\kappa(G)$ is the size of a minimal vertex cut for $G$
- $\lambda(G)$ is the size of a minimal edge cut

## Theorem

$\kappa(G) \leq \lambda(G) \leq \min_{v \in V(G)}\{\delta(v)\}$

# $\kappa(G) \leq \lambda(G) \leq \min_{v \in V(G)}\{\delta(v)\}$

$\lambda(G) \leq \min_{v \in V(G)}\{\delta(v)\}$ Let $u$ have minimal degree $\Rightarrow$ remove the edges incident with it and $u$ becomes isolated.

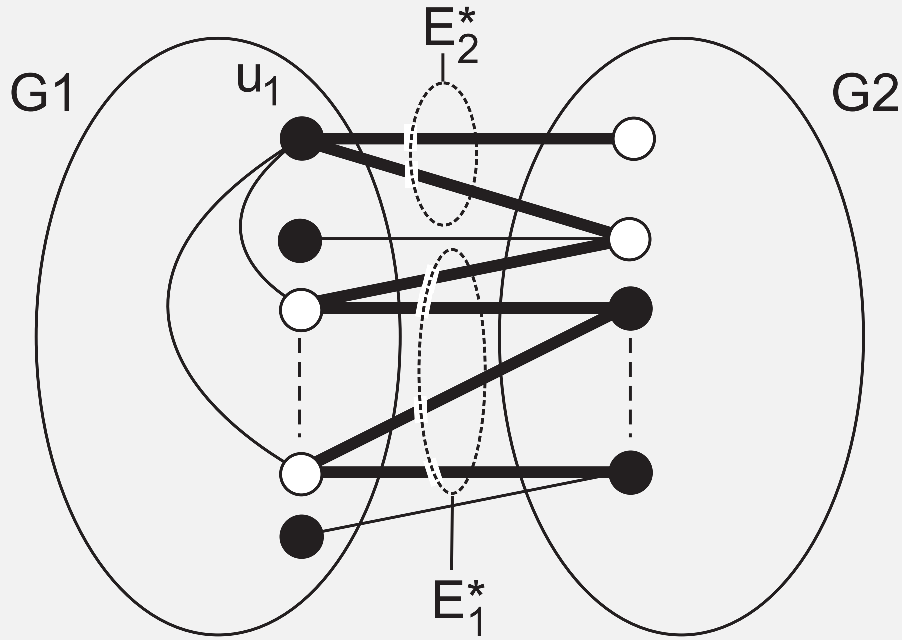
$\kappa(G) \leq \lambda(G)$ Let $E^* = \{\langle u_1, v_1\rangle, \langle u_2, v_2\rangle, \ldots, \langle u_k, v_k\rangle\}$ be an edge cut, with $k = \lambda(G) \Rightarrow G - E^*$ falls into exactly two components $G_1$ and $G_2$ (**why?**).

- Assume there exists $u \in V(G_1) \backslash \{u_1, \ldots, u_k\}$. This means that $\{u_1, \ldots, u_k\}$ is a vertex cut $\Rightarrow \kappa(G) \leq k$.

# $\kappa(G) \le \lambda(G) \le \min_{v\in V(G)}\{\delta(v)\}$ (cnt'd)

Otherwise, assume $V(G_1) = \{u_1, \ldots, u_k\}$ and consider vertex $u_1$.

- $u_1$ is adjacent to $d_1$ vertices $N_1(u_1)$ from $V(G_1)$ and $d_2$ vertices $N_2(u_1)$ from $V(G_2)$.
- Each $u_i \in N_1(u_1)$ is adjacent to a vertex from $V(G_2)$.
- Let $E_1^* = \{\langle u,v\rangle \in E^* | u \in N_1(u_1), v \in V(G_2)\}$ $E_2^* = \{\langle u_1,v\rangle \in E^* | v \in N_2(u_1)\}$
- $d_1 + d_2 \le |E_1^*| + d_2 \le |E_1^*| + |E_2^*| \le |E^*| = \lambda(G)$.
- $N_1(u_1) \cup N_2(u_1)$ is a vertex cut with $d_1 + d_2$ vertices.



# What does it take to be connected?

**Definition**

If $\kappa(G) \geq k$ for some $k$, then $G$ is called k-connected.

**Note**

$G$ is k-connected $\Rightarrow \forall v : \delta(v) \geq k$

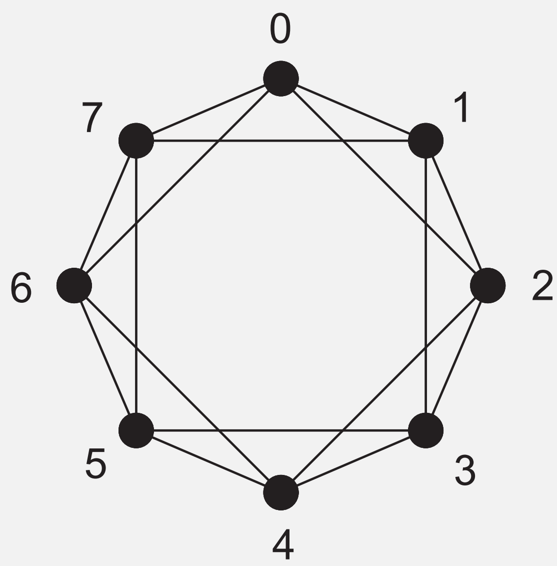
**Issue**

Can we construct a k-connected graph $H_{k,n}$ with $n$ vertices and a minimal number of edges?

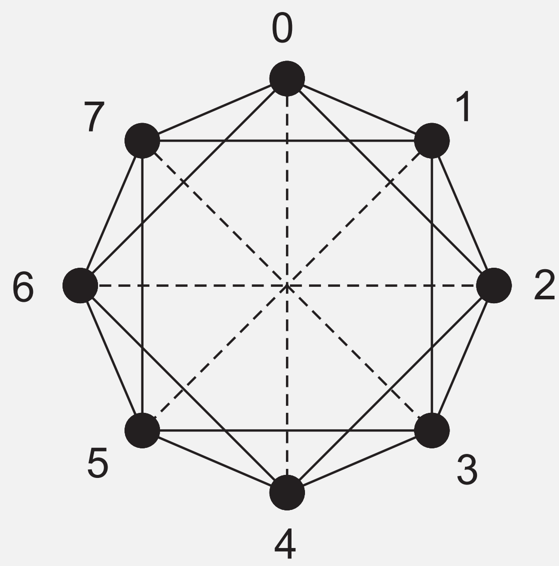
# Harary graphs

**k is even:** Organize vertices $V = \{0, 1, \ldots, n-1\}$ into a "circle."
Connect vertex $i$ to its $k/2$ left-hand (clockwise) neighbors and to its $k/2$ right-hand (counter clockwise) neighbors.

**k is odd, n is even:** Construct $H_{k-1,n}$ and add edges
$\langle 0, \frac{n}{2}\rangle, \langle 1, 1+\frac{n}{2}\rangle, \ldots, \langle \frac{n-2}{2}, n-1\rangle$.



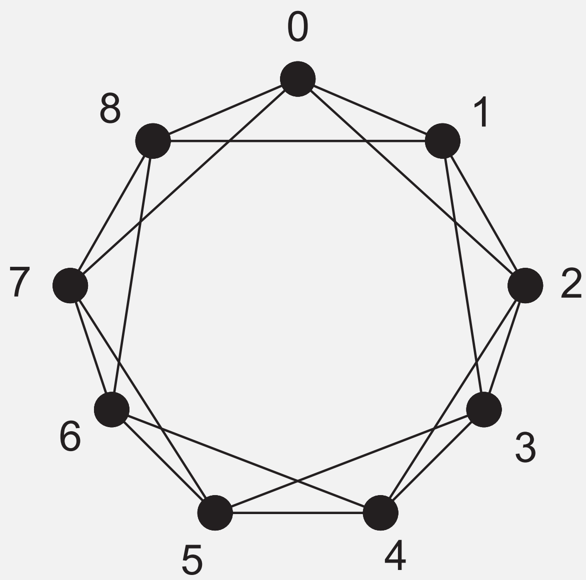

$H_{4,8}$



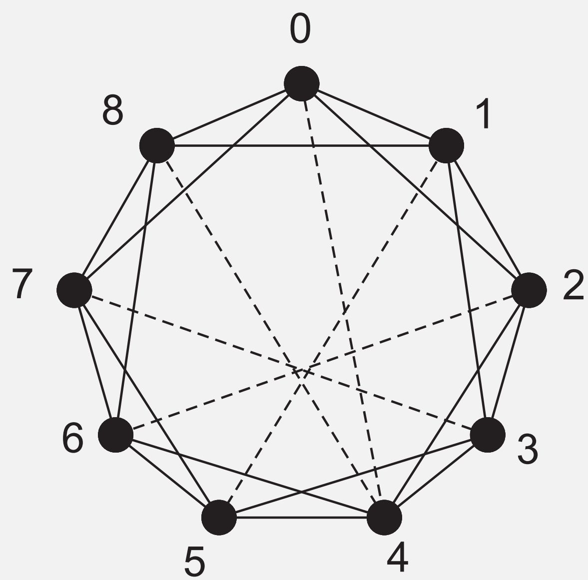

$H_{5,8}$

# Harary graphs

k is odd, n is odd: Construct $H_{k-1,n}$ and add edges
$\langle 0, \frac{n-1}{2} \rangle, \langle 1, 1 + \frac{n-1}{2} \rangle, \ldots, \langle \frac{n-1}{2}, n-1 \rangle$.




$H_{4,9}$




$H_{5,9}$

# Menger's theorem

## Definition

Let $\mathscr{P}(u, v)$ be a collection of paths between vertices $u$ and $v$.
Vertex independent: $\forall P, Q \in \mathscr{P}(u, v) : V(P) \cap V(Q) = \{u, v\}$.
Edge independent: $\forall P, Q \in \mathscr{P}(u, v) : E(P) \cap E(Q) = \emptyset$.

## Theorem (Menger)

*Let $G$ be a graph with two nonadjacent vertices $u$ and $v$. The minimum number of vertices in a vertex cut that disconnects $u$ and $v$ is equal to the maximum number of pairwise vertex-independent paths between $u$ to $v$. The minimum number of edges in an edge cut that disconnects $u$ and $v$, is equal to the maximum number of pairwise edge-independent paths betweeen $u$ and $v$.*

# Menger's theorem

## Mathematical language

Menger's theorem should be read carefully: it mentions pairwise independent paths. In this case, the adjective pairwise is used to make clear that we should always consider pairs of paths when considering independence. And indeed, this makes sense when you would consider trying to count the number of independent paths: being an independent path can only be relative to another path.

To complete the story, also note that the theorem is all about counting the number of $(u, v)$-paths, and not the number of pairs of such paths. In other words, pairwise is an adjective to independent, and not to paths.

# Corollaries

## Corollary

- *$G$ is $k$-connected iff any two distinct vertices are connected by at least $k$ vertex-independent paths.*
- *$G$ is $k$-edge connected iff any two distinct vertices are connected by at least $k$ edge-independent paths.*

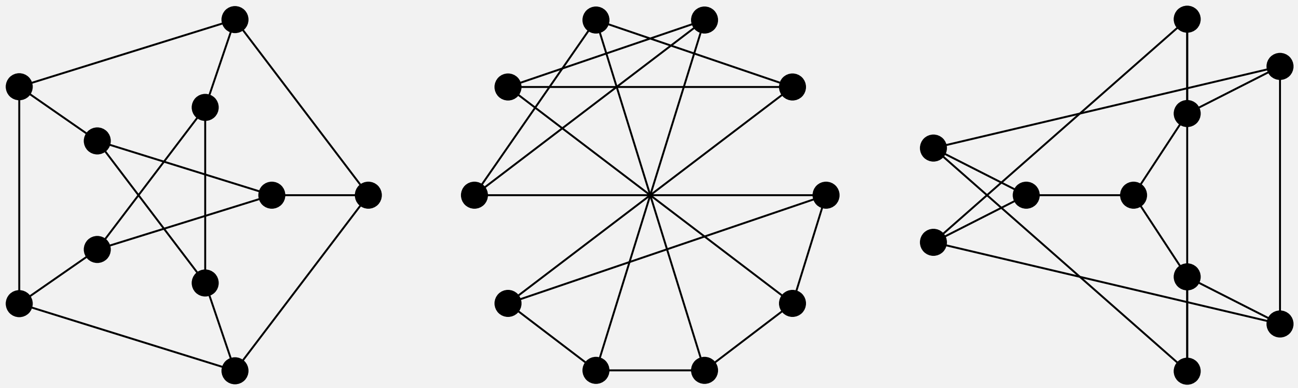
## Corollary

*Each edge of a 2-edge-connected graph lies on a cycle.*

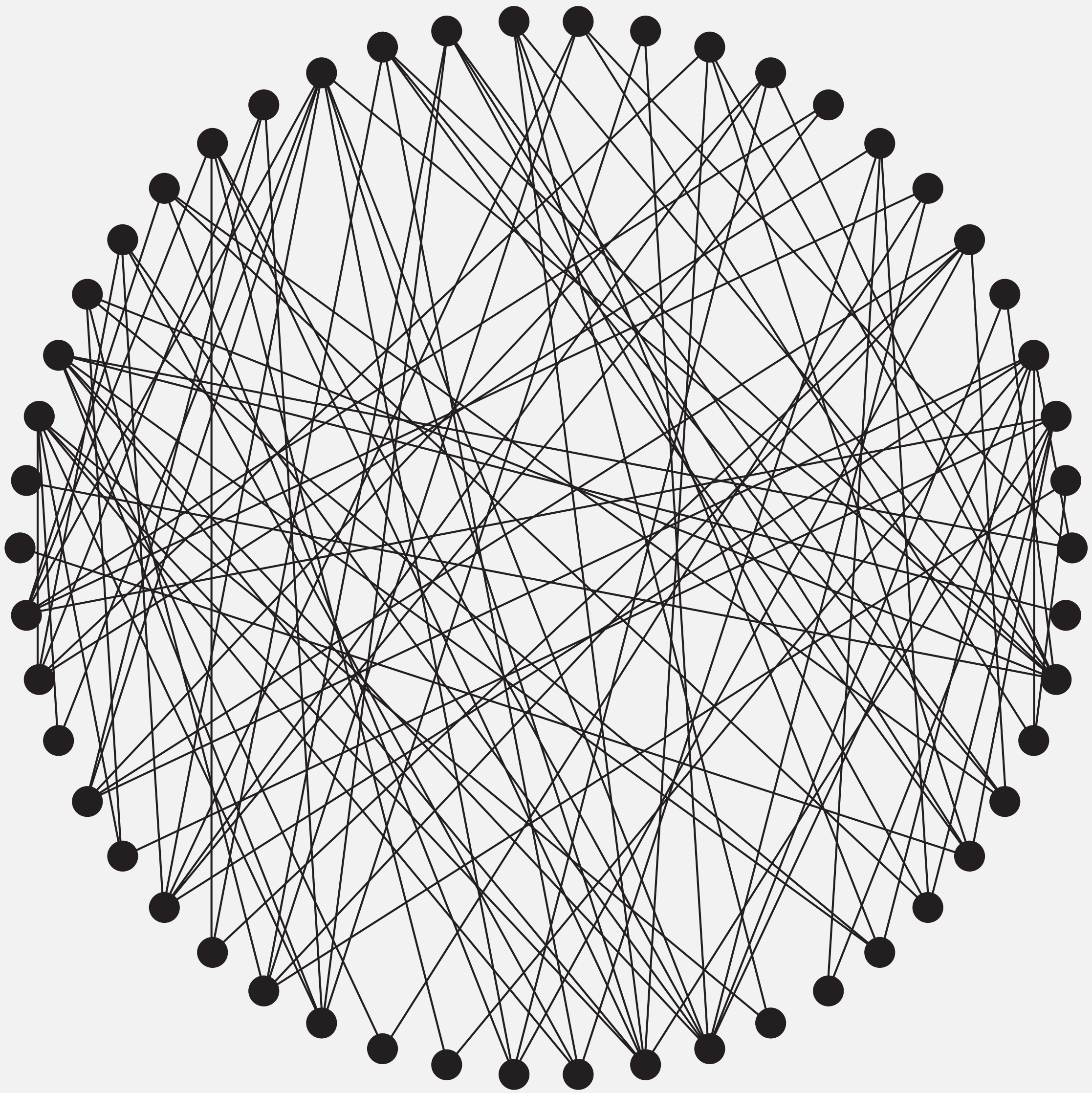
# Drawing graphs

## Observation

It is important to see how you draw a graph, that is, to consider its **graph embedding**.



# Circular embedding



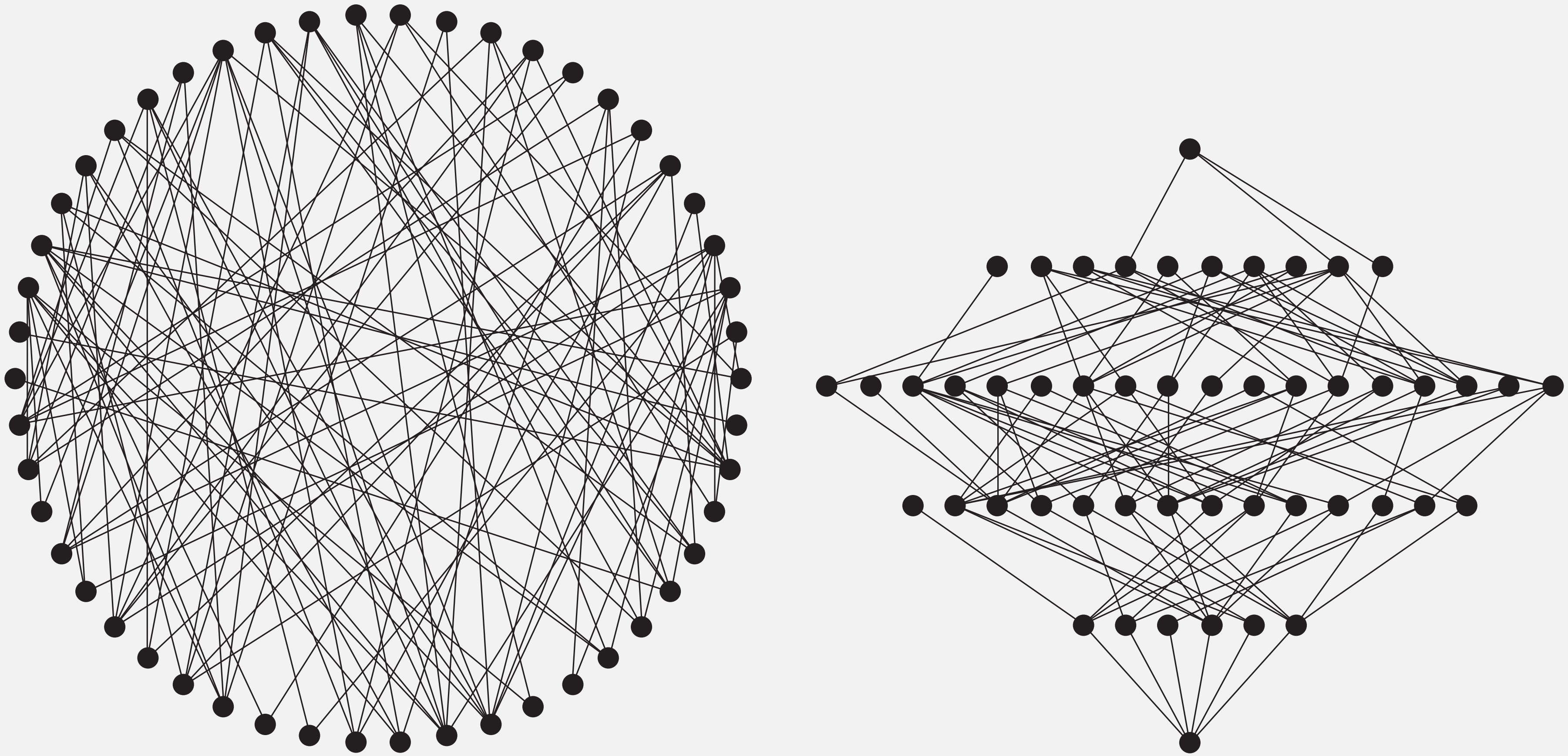
# Ranked embedding

**Definition**

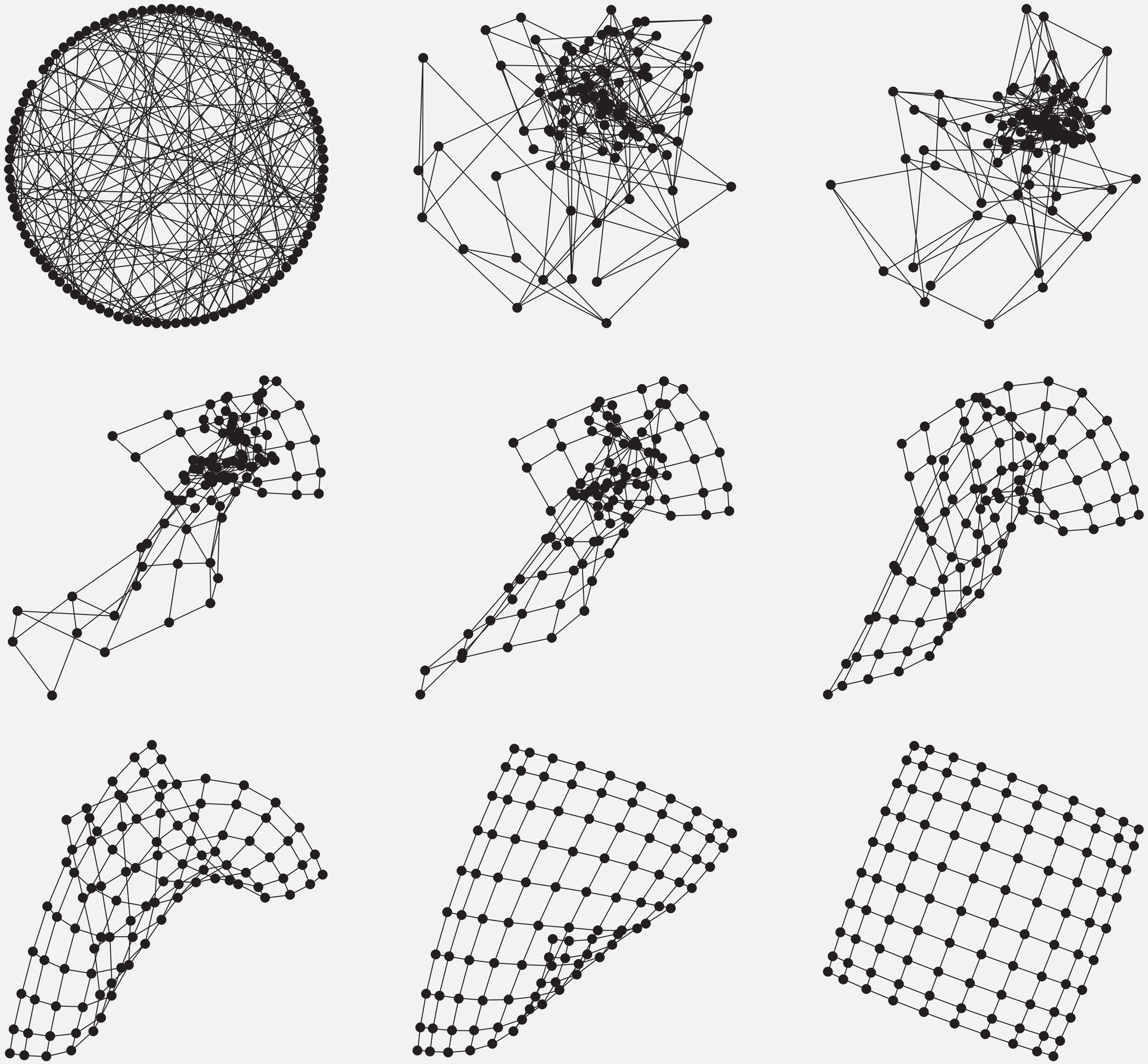
$G$ is bipartite if $V(G) = V_1 \cup V_2$ and $V_1 \cap V_2 = \emptyset$ such that $E(G) \subseteq \{\langle u_1, u_2 \rangle | u_1 \in V_1, u_2 \in V_2\}$.

1. Consider bipartite graph $G$ and vertex $v \in V(G)$
2. Let $N_0^*(v) = \{v\}$
3. Let $N_k^*(v) = N_{k-1}^*(v) \cup \{x \in N(y) | y \in N_{k-1}^*(v)\}, k \geq 1$
4. $N_k(v) = N_k^*(v) - N_{k-1}^*(v)$
5. Draw vertices from $N_k(v)$ on the same vertical line, and vertices from $N_{k-1}(v)$ below (or above) those of $N_k(v)$.

# Ranked embedding



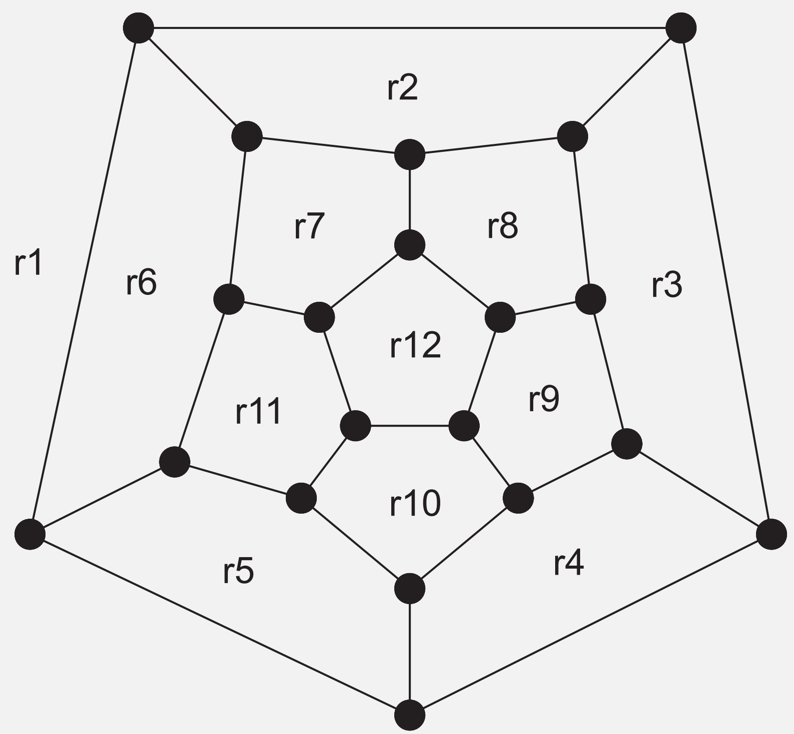
# Spring embedding



# Planar graphs

## Definition

A graph is planar if there exists an embedding in the 2D plane such that no two edges cross. A plane graph is a drawing of a planar graph such that no two edges intersect.




## Theorem (Euler's formula)

*For a plane graph with $n$ vertices, $m$ edges, and $r$ regions: $n - m + r = 2$.*

# Planar graphs: properties

**Theorem**

*For any connected simple planar graph with $n \geq 3$ vertices and $m$ edges: $m \leq 3n-6$*

- Consider region $f$ in a plane graph of $G$
- $\forall$ interior regions: $B(f)$ denotes number of edges enclosing $f$.
  Note: $B(f) \geq 3$.
- $n \geq 3 \Rightarrow$ exterior region bounded by at least 3 edges.
- $r$ regions $\Rightarrow \sum B(f) \geq 3r$
- $\sum B(f)$ counts edges once or twice $\Rightarrow \sum B(f) \leq 2m$
- $3r \leq \sum B(f) \leq 2m \Rightarrow m = n + r - 2 \leq n + \frac{2}{3}m - 2 \Rightarrow$
  $m \leq 3n-6$