STRUTTURE RELAZIONALI, GRAFI E ORDINAMENTI

(parte 3)

Stefania Bandini

GRAFI

GRAFI

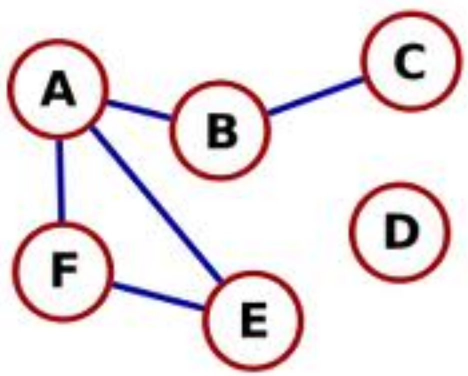
I grafi sono strutture matematiche discrete che rivestono interesse sia per la matematica che per un'ampia gamma di campi applicativi.

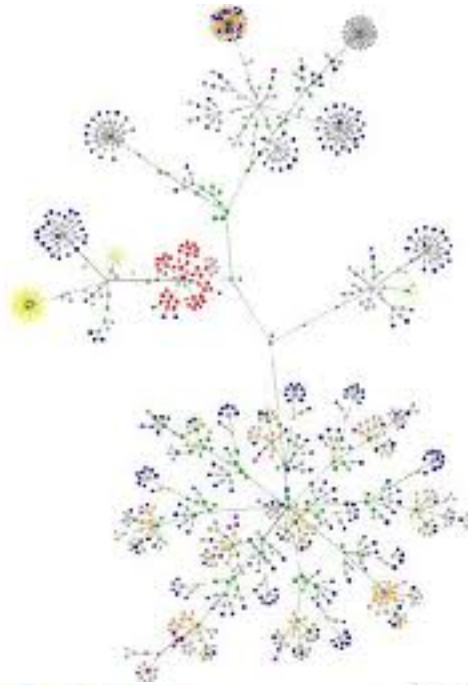
In ambito matematico il loro studio, la teoria dei grafi, costituisce un'importante parte della combinatoria; i grafi inoltre sono utilizzati in aree come topologia, teoria degli automi, funzioni speciali, geometria dei poliedri, etc.

I grafi si incontrano in vari capitoli dell'informatica (per schematizzare programmi, circuiti, reti di computer, mappe di siti).

Sono alla base di modelli di sistemi e processi studiati nell'ingegneria, nella chimica, nella biologia molecolare, nella ricerca operativa, nella organizzazione aziendale, nella geografia (sistemi fluviali, reti stradali, trasporti), nella linguistica strutturale, nella storia (alberi genealogici, filologia dei testi).

FONDAMENTI DELL'INFORMATICA





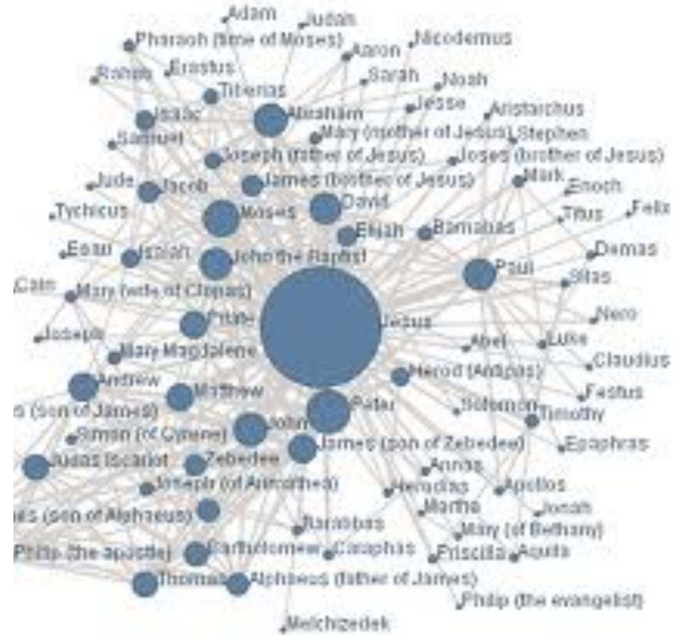
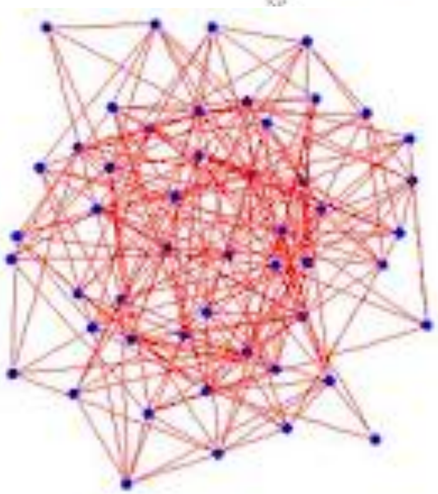


Figura 10



GRAFI

In matematica, in informatica e, più in particolare, in geometria combinatoria, i grafi sono oggetti discreti che permettono di schematizzare una grande varietà di situazioni e di processi e spesso di consentirne l'analisi in termini quantitativi e algoritmici.
In termini informali, per grafo si intende una struttura costituita da:

oggetti semplici, detti vertici (*vertices*) o nodi (*nodes*),
collegamenti tra i vertici. I collegamenti possono essere:

- **orientati**, e in questo caso sono detti archi (*arcs*), e il grafo è detto **orientato**
- **non orientati**, e in questo caso sono detti spigoli (*edges*), e il grafo è detto **non orientato**
- eventualmente **dati associati a nodi e/o collegamenti**

GRAFI

Un grafo viene generalmente raffigurato sul piano da punti o cerchietti, che rappresentano i nodi; archi o spigoli sono rappresentati da segmenti o curve che collegano due nodi. In questo caso, il posizionamento dei nodi e la forma degli archi o spigoli è irrilevante, dal momento che a contare sono solo i nodi e le relazioni tra essi. In altri termini, lo stesso grafo può essere disegnato in molti modi diversi senza modificarne le proprietà.

GRAFI

Le strutture che possono essere rappresentate da grafi sono onnipresenti e molti problemi di interesse pratico possono essere formulati come questioni relative a grafi. In particolare, le reti possono essere descritte in forma di grafi. Ad esempio, la struttura dei link della Wikipedia, come tutti gli ipertesti, può essere rappresentata da un grafo orientato, dove i vertici sono gli articoli e gli archi rappresentato l'esistenza di un link tra un articolo e l'altro. I grafi orientati sono anche utilizzati per rappresentare le macchine a stati finiti e molti altri formalismi, come ad esempio diagrammi di flusso, catene di Markov, schemi entità-relazione, reti di Petri e molti altri.

Lo sviluppo di algoritmi per manipolare i grafi è una delle aree di maggior interesse dell'informatica.

GRAFI

Nel caso una relazione binaria G sia definita su un insieme V: $G \subseteq V \times V$ la relazione binaria può essere rappresentata mediante un *grafo orientato* (talvolta detto *diretto* o anche *digrafo*, dall'inglese "directed graph").

Gli elementi di V sono detti *vertici* o *nodi* del grafo e gli elementi di G sono detti *archi*.

Rappresentazioni di una relazioni binaria G

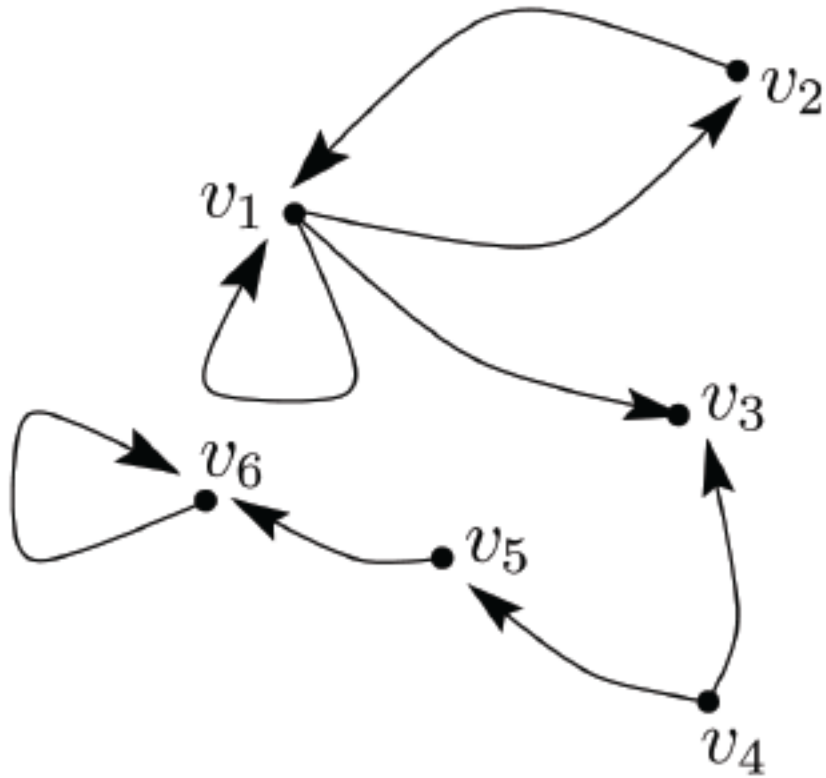
Sia dato un insieme di vertici $V = \{v_1, \ v_2, \ v_3, \ v_4, \ v_5, \ v_6\}$. Sia data la relazione binaria G su $V \times V$

$$\{\langle v_1, v_1\rangle, \langle v_1, v_2\rangle, \langle v_1, v_3\rangle, \langle v_2, v_1\rangle, \langle v_4, v_3\rangle, \langle v_4, v_5\rangle, \langle v_5, v_6\rangle, \langle v_6, v_6\rangle\}$$

La seguente tabella rappresenta la relazione binaria G

v_1	v_1
v_1	v_2
v_1	v_3
v_2	v_1
v_4	v_3
v_4	v_5
v_5	v_6
v_6	v_6

Rappresentazioni di una relazioni binaria G



Grafo associato alla relazione introdotta nell'esempio precedente

Rappresentazioni di relazioni

La relazione G introdotta nell'esempio precedente può anche essere rappresentata mediante la seguente matrice booleana

$$\begin{vmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

TERMINOLOGIA

Un arco che congiunge v_i a v_j si dice *uscente* da v_i ed *entrante* in v_j.

Il numero di archi uscenti da un nodo è detto *grado di uscita* del nodo.

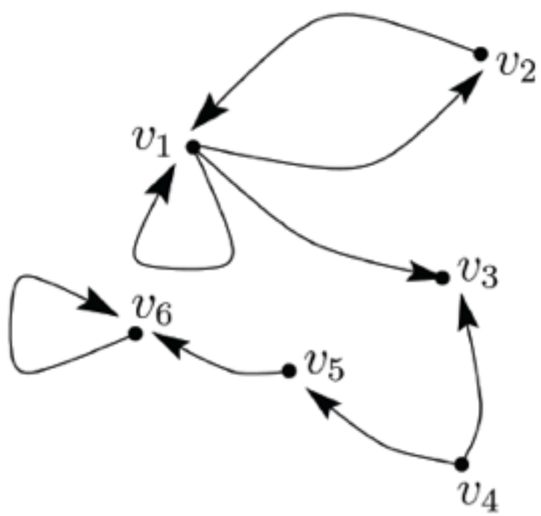
Il numero di archi entranti in un nodo è detto *grado di ingresso* del nodo.

Un nodo di un grafo si chiama *nodo sorgente* se non ha archi entranti, *nodo pozzo* se non ha archi uscenti.

Un nodo di un grafo si dice *nodo isolato* se non ha archi entranti né uscenti.

PROPRIETA’ DI NODI

Esempio



v_1 ha grado di ingresso 2 e uscita 3;

v_2 ha grado di ingresso 1 e uscita 1;

v_3 ha grado di ingresso 2 e uscita 0 (nodo pozzo);

v_4 ha grado di ingresso 0 (nodo sorgente) e uscita 2;

v_5 ha grado di ingresso 1 e uscita 1;

v_6 ha grado di ingresso 2 e uscita 1.

TERMINOLOGIA

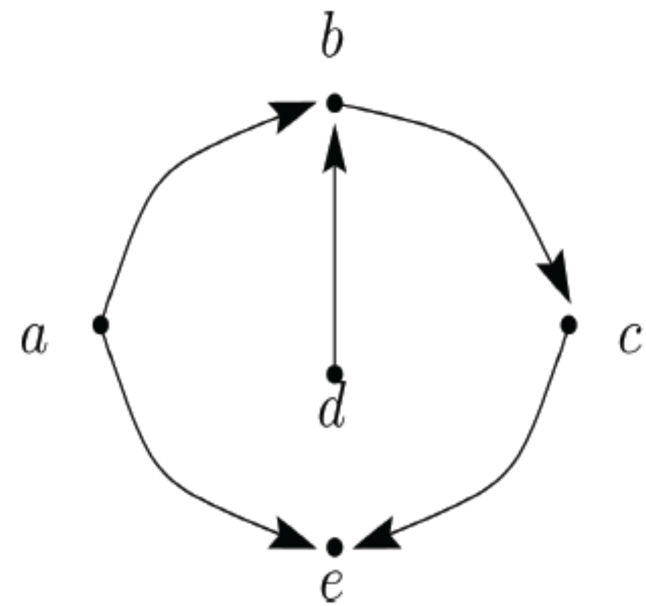
Un *cammino* tra due nodi v_{in} e v_{fin} di un grafo è una sequenza finita di nodi $\langle v_1, v_2, \ldots v_n \rangle$ con $v_1 = v_{in}$ e $v_n = v_{fin}$, dove ciascun nodo è collegato al successivo della sequenza da un arco uscente dal primo ed entrante nel secondo.

Un *semicammino* tra due nodi v_{in} e v_{fin} di un grafo è una sequenza finita $\langle v_1, v_2, \ldots v_n \rangle$ con $v_1 = v_{in}$ e $v_n = v_{fin}$, dove ciascun nodo è collegato al successivo della sequenza da un arco di direzione arbitraria.

La *lunghezza* di un cammino (o di un semicammino) tra due nodi di un grafo è data dal numero di archi che lo compongono; quindi essa è uguale al numero dei nodi della sequenza, meno 1.

CAMMINI

Nel grafo in figura ci sono due **cammini** tra *a* ed *e*



- un cammino di lunghezza 1: $\langle a, e \rangle$;
- un cammino di lunghezza 3: $\langle a, b, c, e \rangle$.

Non ci sono cammini tra a e d, ma c'è un semicammino di lunghezza 2: $\langle a, b, d \rangle$, un semicammino di lunghezza 4: $\langle a, e, c, b, d \rangle$, uno di lunghezza 6: $\langle a, b, c, e, a, b, d \rangle$, eccetera.

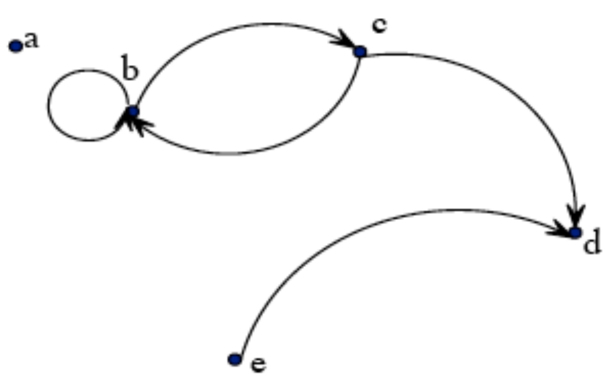
TERMINOLOGIA

Un grafo si dice *connesso* se, dati due nodi qualunque purché distinti, esiste sempre un semicammino tra essi.

Un *ciclo* intorno ad un nodo v di un grafo è un cammino in cui $v = v_{in} = v_{fin}$.

Un *semiciclo* intorno ad un nodo di un grafo è un semicammino in cui $v = v_{in} = v_{fin}$.

Un *cappio* intorno ad un nodo è un cammino di lunghezza 1.



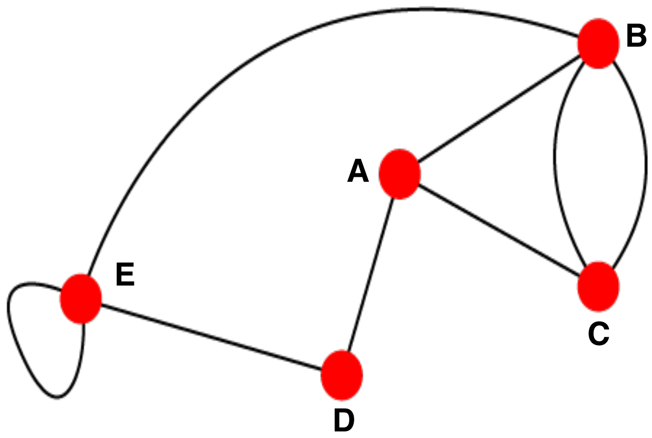
Esempio di grafo non connesso.

PATHS

A *path is* a sequence of nodes in which each node is adjacent to the next one

$P_{i0,in}$ of length *n* between nodes i_0 and i_n is an ordered collection of *n+1* nodes and *n* links

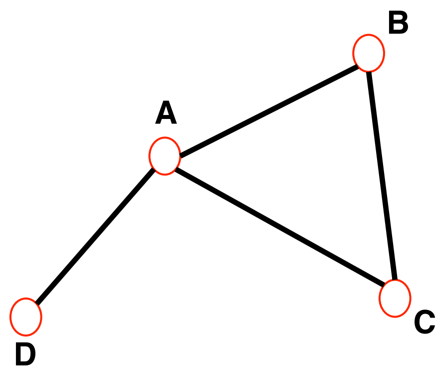
$$P_n = \{i_0, i_1, i_2, \ldots, i_n\} \qquad P_n = \{(i_0, i_1), (i_1, i_2), (i_2, i_3), \ldots, (i_{n-1}, i_n)\}$$



•A path can intersect itself and pass through the same link repeatedly. Each time a link is crossed, it is counted separately

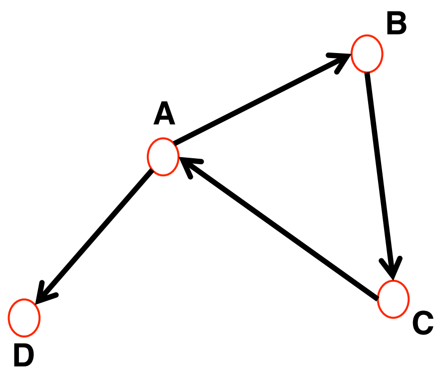
•A legitimate path on the graph on the right:
ABCBCADEEBA

• In a directed network, the path can follow only the direction of an arrow.



The *distance (shortest path, geodesic path)* between two nodes is defined as the number of edges along the shortest path connecting them.

*If the two nodes are disconnected, the distance is infinity.



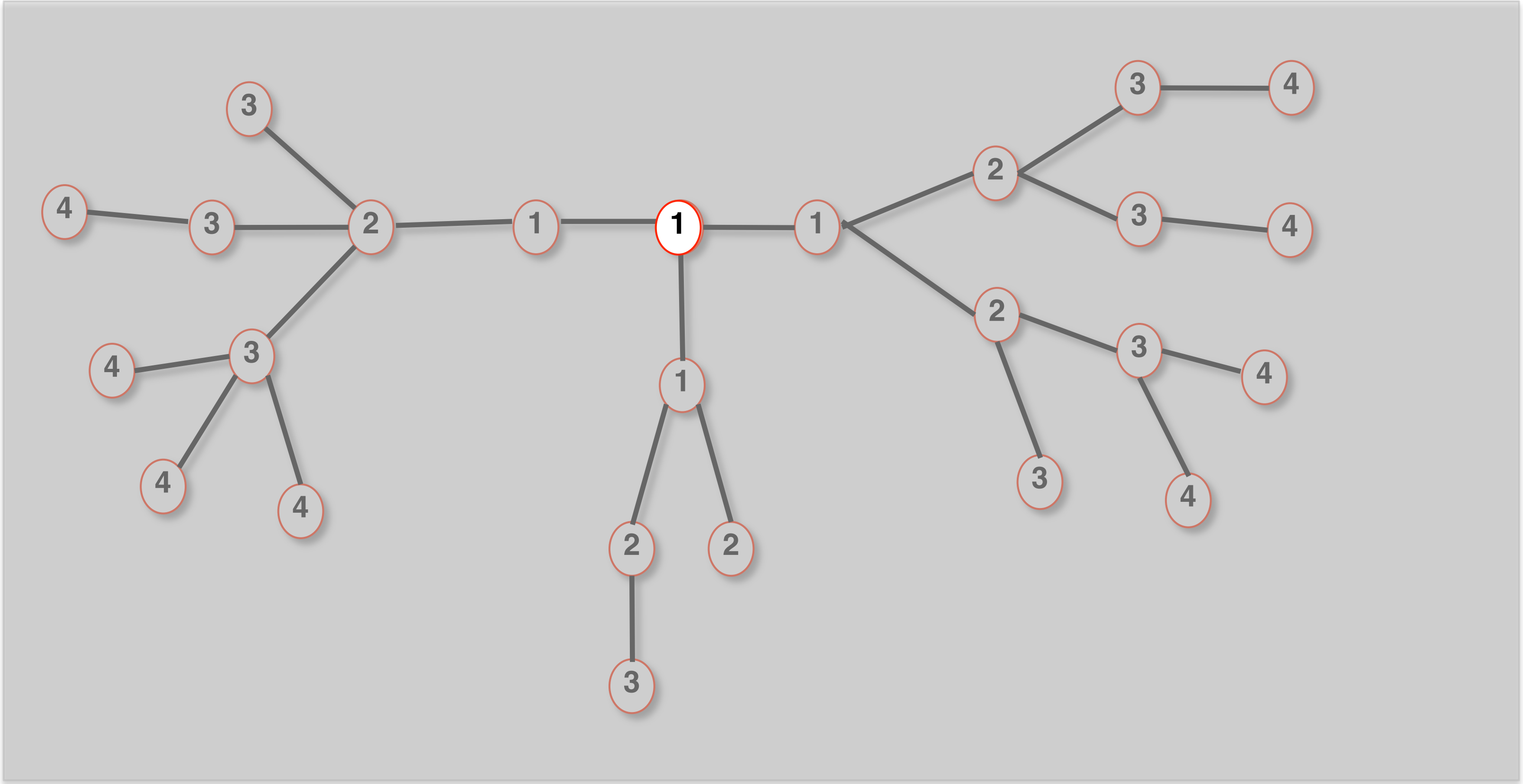
In directed graphs each path needs to follow the direction of the arrows.

Thus in a digraph the distance from node A to B (on an AB path) is generally different from the distance from node B to A (on a BCA path).

FINDING DISTANCES: BREADTH FIRST SEATCH

Distance between node 1 and node 4:

1.Start at 1.

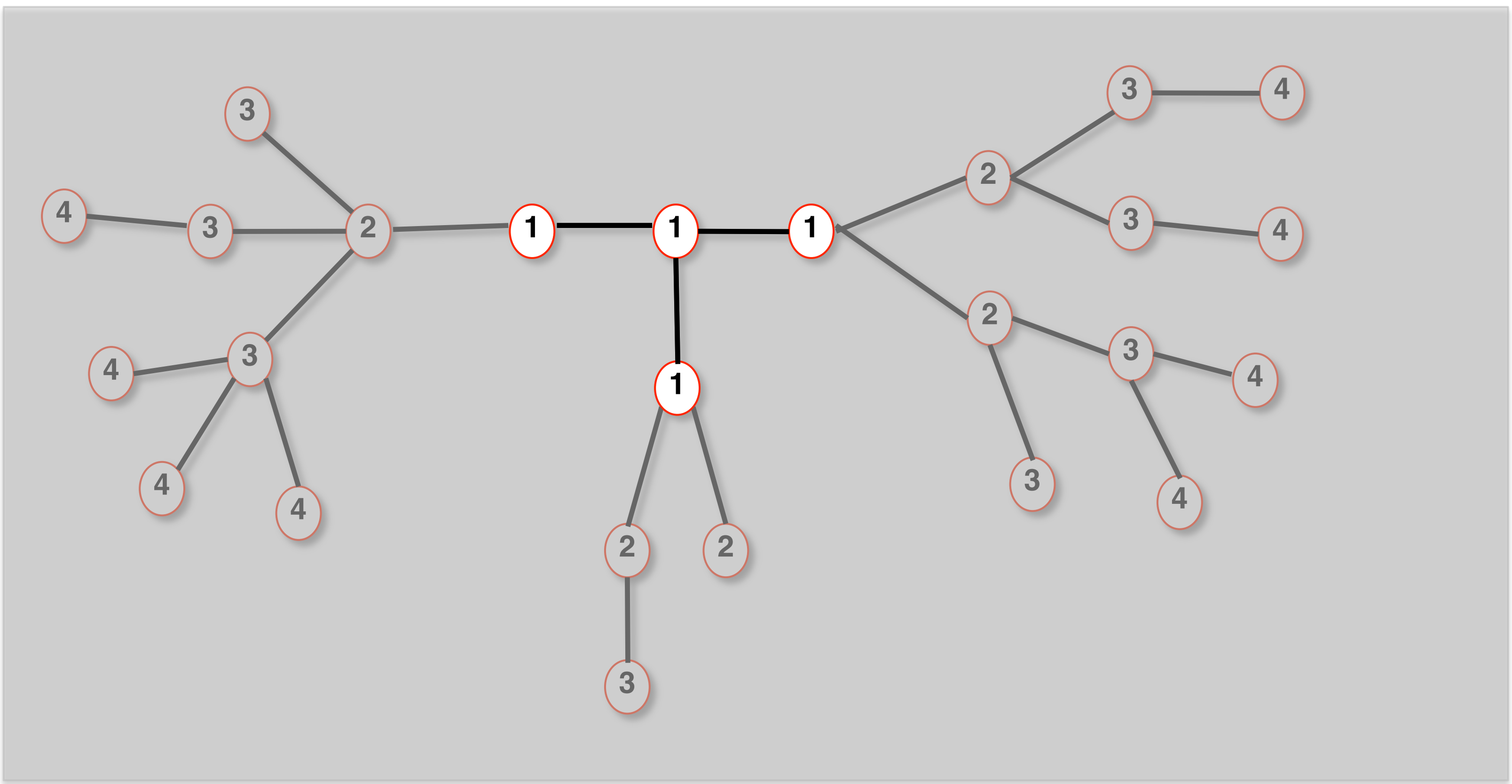


FINDING DISTANCES: BREADTH FIRST SEATCH

Distance between node 1 and node 4:

1.Start at 1.

2.Find the nodes adjacent to 1. Mark them as at distance 1. Put them in a queue.



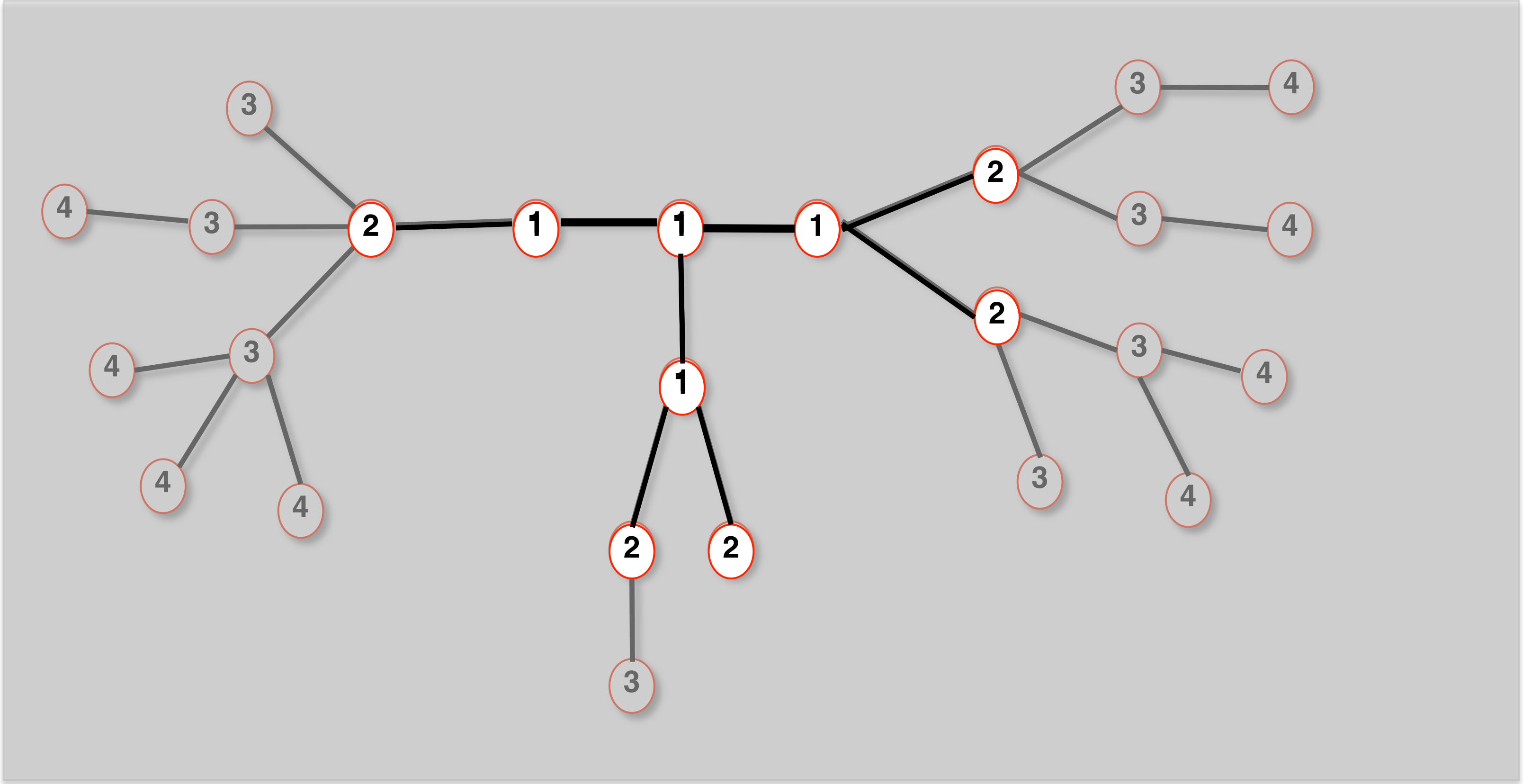
FINDING DISTANCES: BREADTH FIRST SEATCH

Distance between node 1 and node 4:

1.Start at 1.

2.Find the nodes adjacent to 1. Mark them as at distance 1. Put them in a queue.

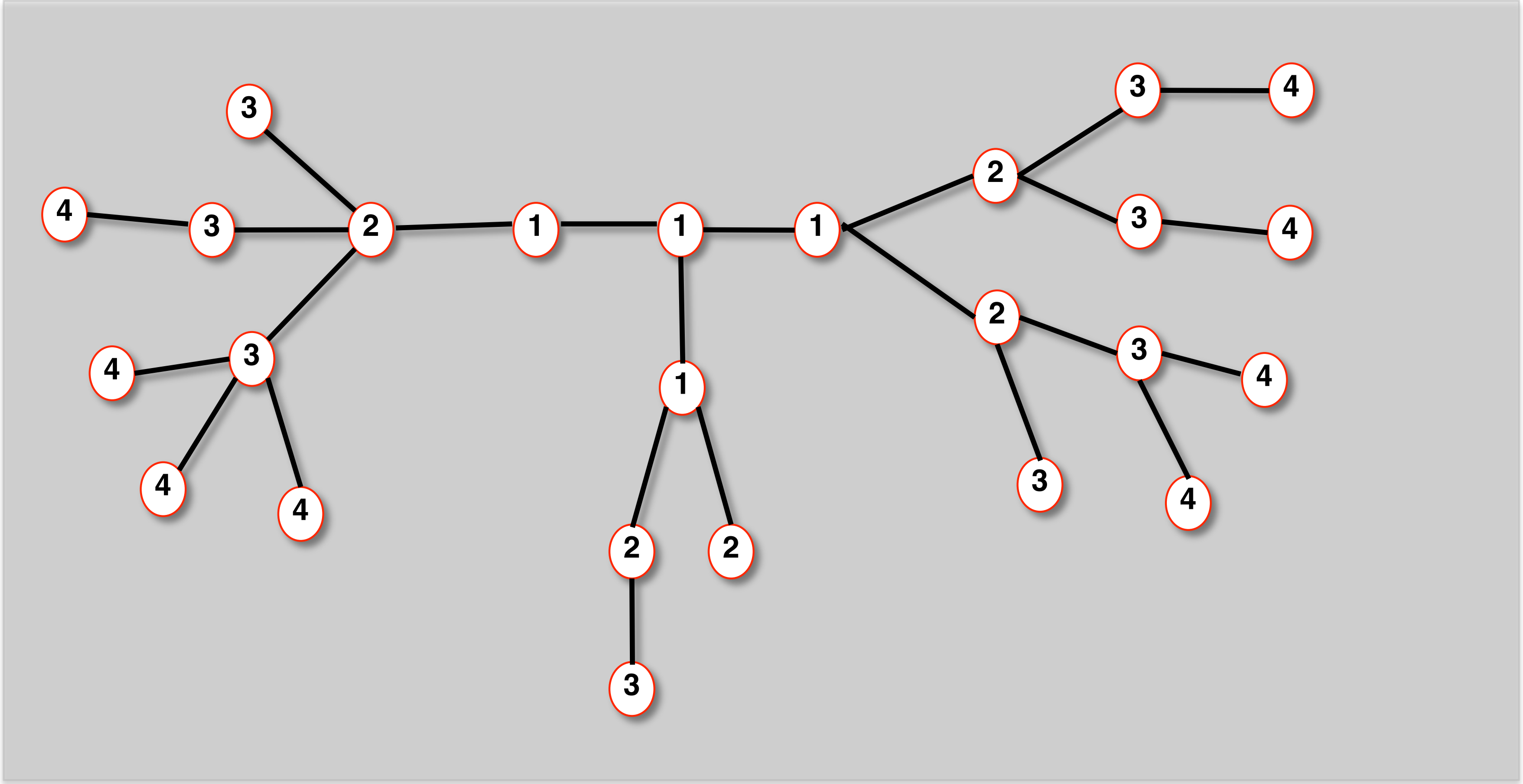
3.Take the first node out of the queue. Find the unmarked nodes adjacent to it in the graph. Mark them with the label of 2. Put them in the queue.



FINDING DISTANCES: BREADTH FIRST SEATCH

Distance between node 1 and node 4:

1. Repeat until you find node 4 or there are no more nodes in the queue.
2. The distance between 1 and 4 is the label of 4 or, if 4 does not have a label, infinity.

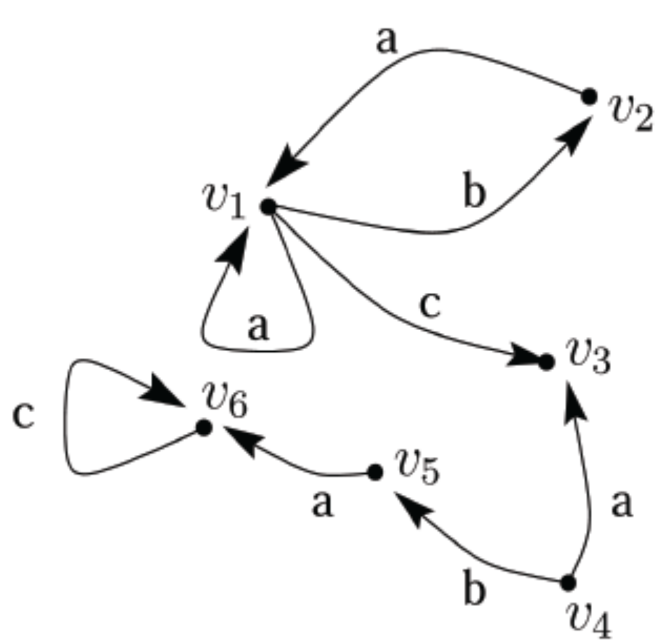


GRAFO ETICHETTATO

Un *grafo etichettato* è una funzione che associa ad ogni arco di un grafo un'etichetta.

Sia $G \subseteq V \times V$ un grafo.

Un grafo etichettato è una funzione $E : G \mapsto N$ dove N è un insieme di nomi, o etichette.



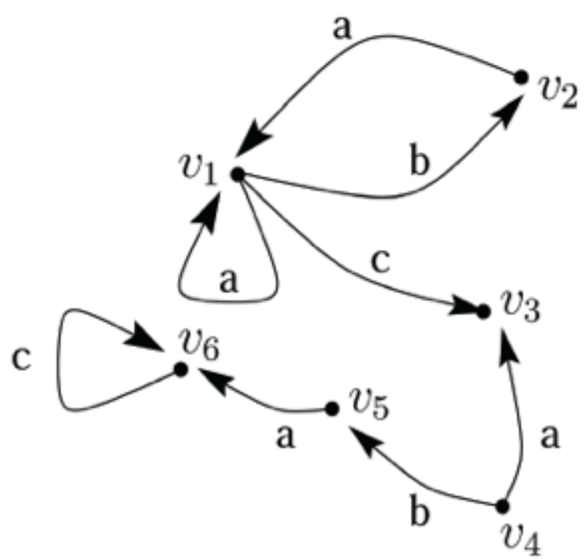
GRAFO ETICHETTATO NEGLI ARCHI

Sia dato l'insieme di vertici V e la relazione binaria G su $V \times V$

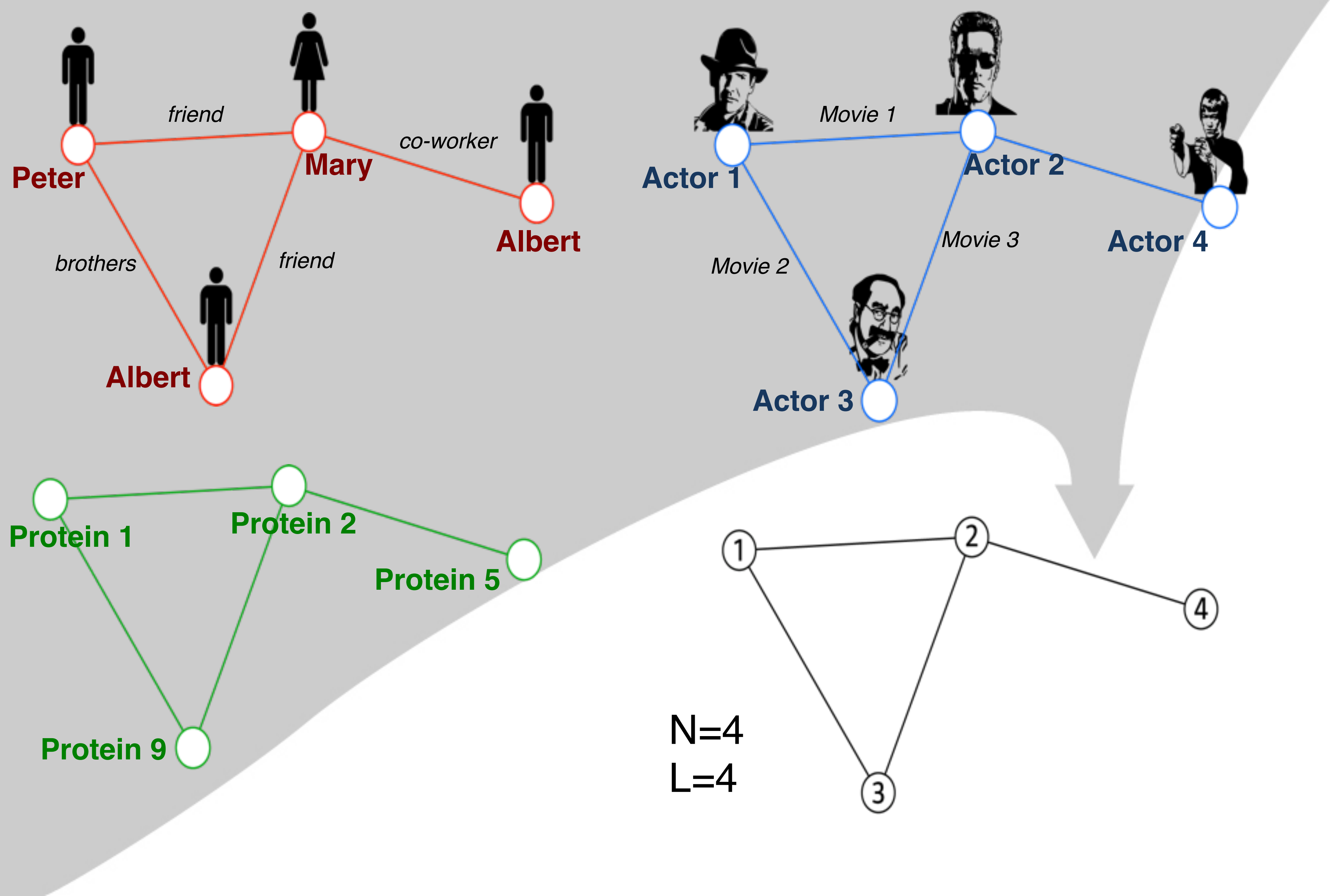
Sia $N = \{a, b, c\}$ un insieme di nomi e sia $E : G \mapsto N$

la funzione di etichettatura di G definita dalla seguente tabella

v_1	v_1	a
v_1	v_2	b
v_1	v_3	c
v_2	v_1	a
v_4	v_3	a
v_4	v_5	b
v_5	v_6	a
v_6	v_6	c



A COMMON LANGUAGE

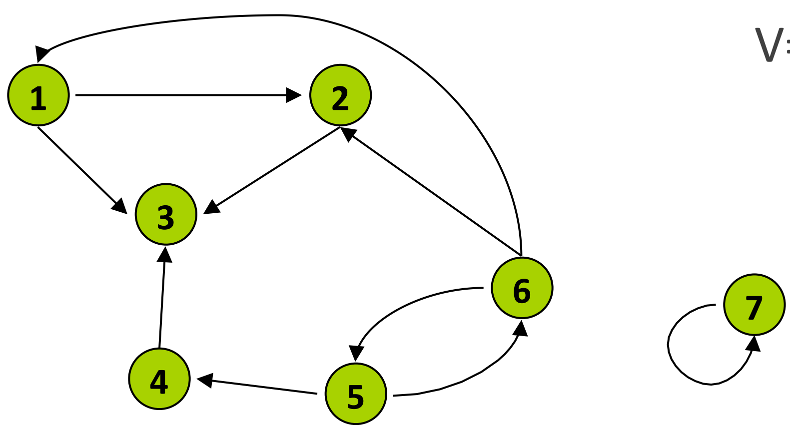


Grafo diretto

Un grafo diretto **G** è definito da due insiemi (V, E).

V è l'insieme dei **vertici** {1, 2, 3, … , n} (detti anche **nodi**).

E è l'insieme degli **archi diretti (u, v)** con u e v in V



V= {1, 2, 3, 4, 5, 6, 7}

Grafo non diretto

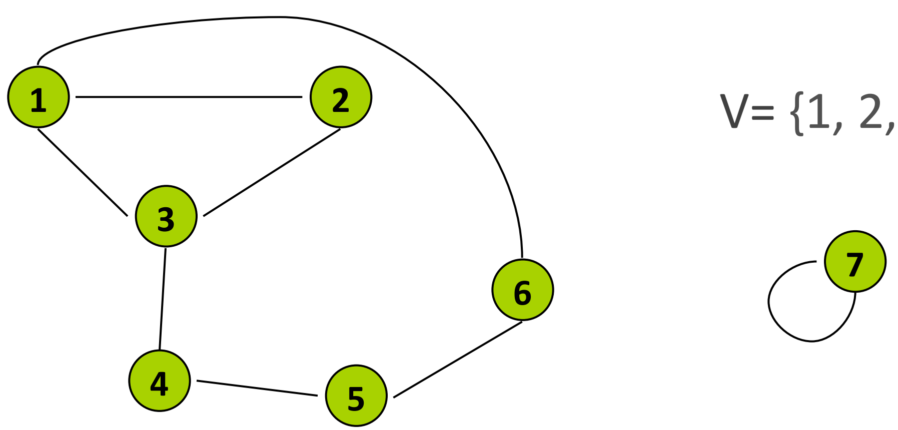
Un grafo non diretto **G'** è definito da due insiemi (V', E').

V' è l'insieme dei **vertici** {1, 2, 3, … , n} (detti anche **nodi**).

E' è l'insieme degli **archi non diretti (u, v)** con u e v in V'.

Nota: *(u, v) risulta identico a (v, u)!*

Quindi, per rappresentare il grafo G' con un grafo diretto, basta sostituire ogni arco non diretto (u,v) con i due archi diretti (u,v) e (v,u).



V= {1, 2, 3, 4, 5, 6, 7}

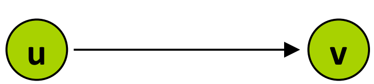
Incidenza e adiacenza

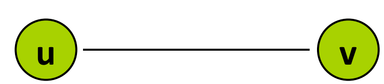
Nel grafo diretto G=(V, E), si dice che l’arco diretto **(u, v)** è **incidente** da **u** a **v** (l’arco esce da u ed entra in v).

Nel grafo non diretto G’=(V’, E’), si dice semplicemente che l’arco non diretto **(u, v)** è **incidente su u** e **v.**

I vertici **u** e **v** risultano **adiacenti** tra di loro

In un grafo diretto G si può scrive **u → v**, per mettere in evidenza il verso di percorrenza dei due vertici adiacenti.



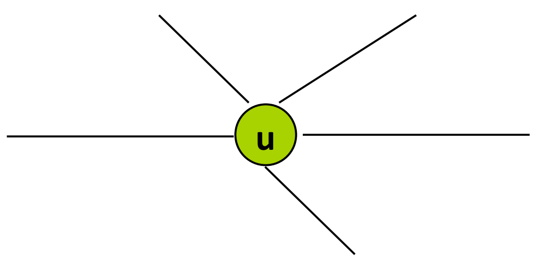


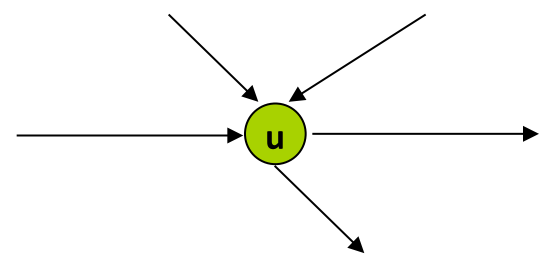
Grado di un nodo

Il grado di un nodo u corrisponde al **numero di archi incidenti** con u.

In un grafo diretto si può anche calcolare il **grado uscente** o il **grado entrante** di u, rispettivamente il numero di archi uscenti da u o entranti in u.

$\delta^-(u) = 3$





$\delta(u) = 5$

$\delta^+(u) = 2$

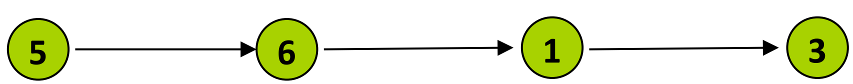
Cammini su un grafo

Un **cammino** è una sequenza $\langle v_1, v_2, \ldots, v_k \rangle$ di nodi a due a due adiacenti, dove (v_i, v_{i+1}) con $i = 1, \ldots, k-1$ è un arco del grafo.

Se $v_1 = u$ e $v_k = v$, allora si dice che esiste un cammino tra u e v (u~>v) e che v è **raggiungibile** da u.

Il **cammino** di dice **semplice** se i vertici sono tutti distinti.

Un **sottocammino** $\langle v_i, v_{i+1}, \ldots, v_j \rangle$ (con $0 \le i \le j \le k$) è una sottosequenza continua dei vertici che compongono un cammino.

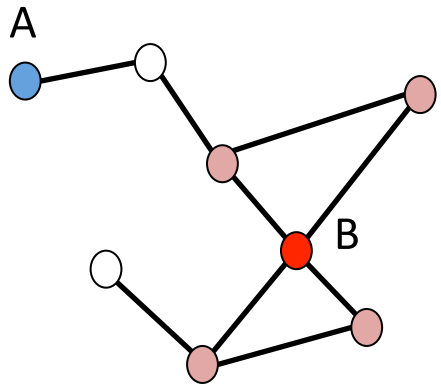


Un cammino semplice del primo grafo diretto:

$\langle 5, 6, 1, 3 \rangle$

NODE DEGREES

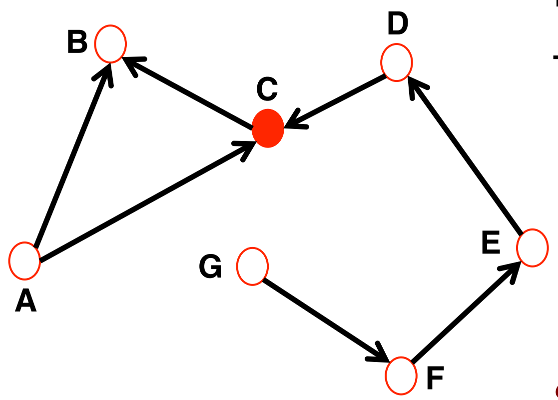
Undirected



Node degree: the number of links connected to the node.

$$k_A = 1 \qquad k_B = 4$$

Directed



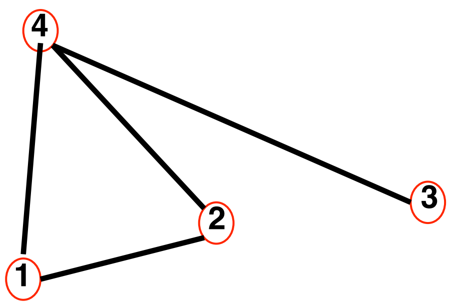
In *directed networks* we can define an in-degree and out-degree.

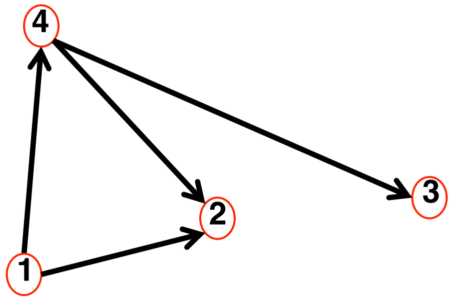
The (total) degree is the sum of in- and out-degree.

$$k_C^{in} = 2 \quad k_C^{out} = 1 \qquad k_C = 3$$

Source: a node with $k^{in} = 0$; Sink: a node with $k^{out} = 0$.

ADJACENCY MATRIX





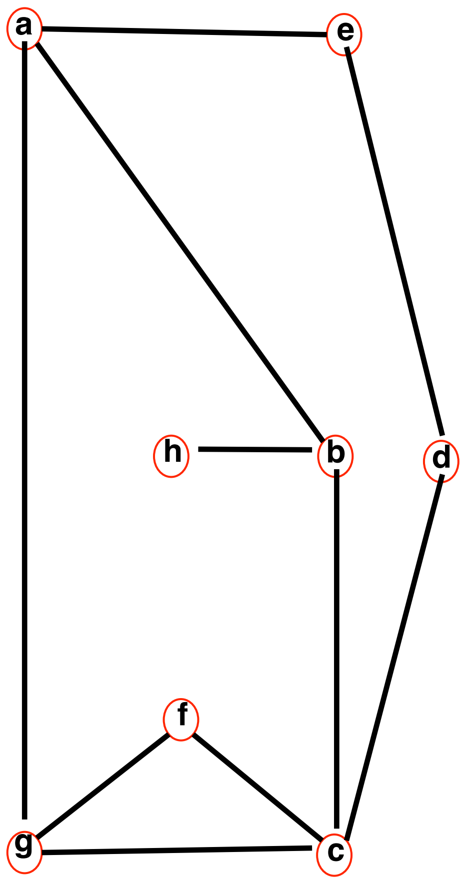
A_{ij}=1 if there is a link between node *i* and *j*

A_{ij}=0 if nodes *i* and *j* are not connected to each other.

$$A_{ij} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \qquad A_{ij} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

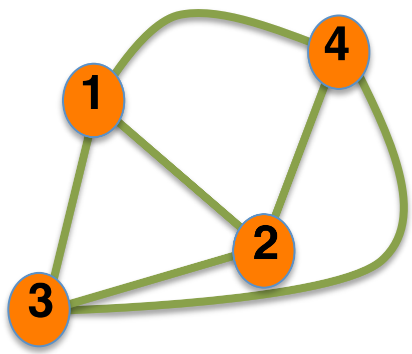
Note that for a directed graph (right) the matrix is not symmetric.

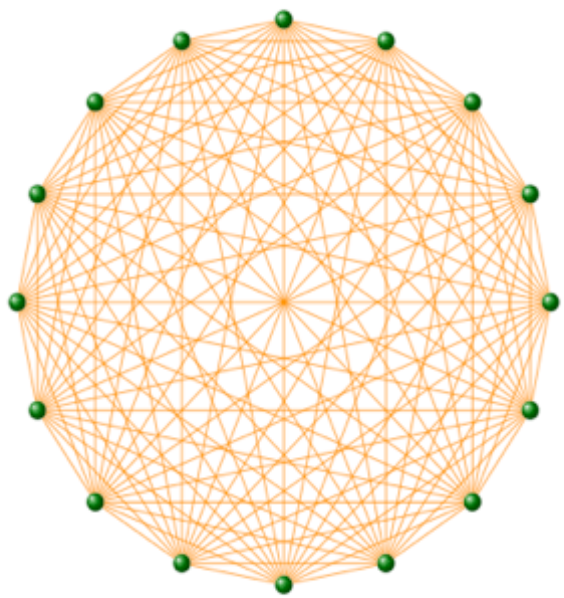
ADJACENCY MATRIX



	a	b	c	d	e	f	g	h
a	0	1	0	0	1	0	1	0
b	1	0	1	0	0	0	0	1
c	0	1	0	1	0	1	1	0
d	0	0	1	0	1	0	0	0
e	1	0	0	1	0	0	0	0
f	0	0	1	0	0	0	1	0
g	1	0	1	0	0	0	0	0
h	0	1	0	0	0	0	0	0

Complete Graph
(undirected)





$$A_{ij} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

$$A_{ii} = 0 \qquad A_{i \neq j} = 1$$

$$L = L_{\max} = \frac{N(N-1)}{2} \qquad \langle k \rangle = N - 1$$

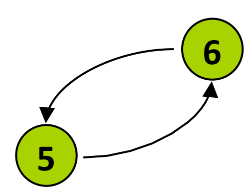
Actor network, protein-protein interactions

Cicli su un grafo

Un **ciclo** è un cammino **$<v_1, v_2, \dots, v_k>$** dove **$v_1 = v_k$**

Quando il sottocammino $<v_1, v_2, \dots, v_{k-1}>$ del ciclo è semplice allora di dice che il **ciclo è semplice**.

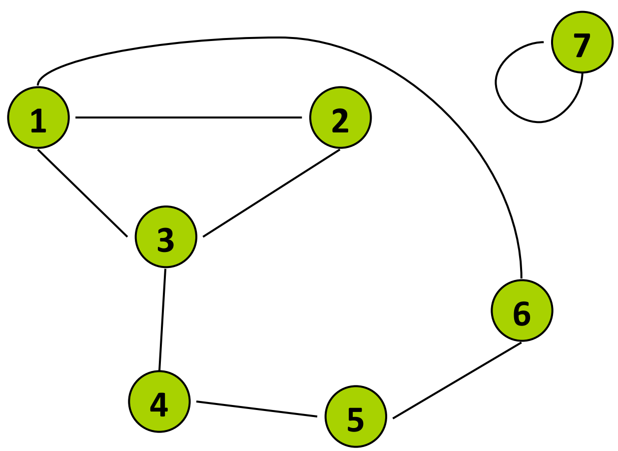
Un grafo senza cicli è detto **aciclico**.



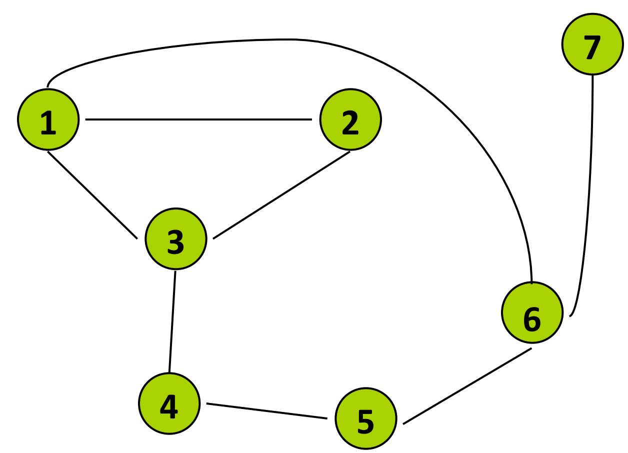
Grafo connesso

Un **grafo non diretto** è **connesso** se ogni coppia di vertici è connesso da un cammino.

Un **grafo diretto** è detto **fortemente connesso** se esiste un cammino tra ogni coppia di vertici. Quindi si ha u~> v e v~> u per ogni coppia di vertici (u, v).



G_1 non connesso

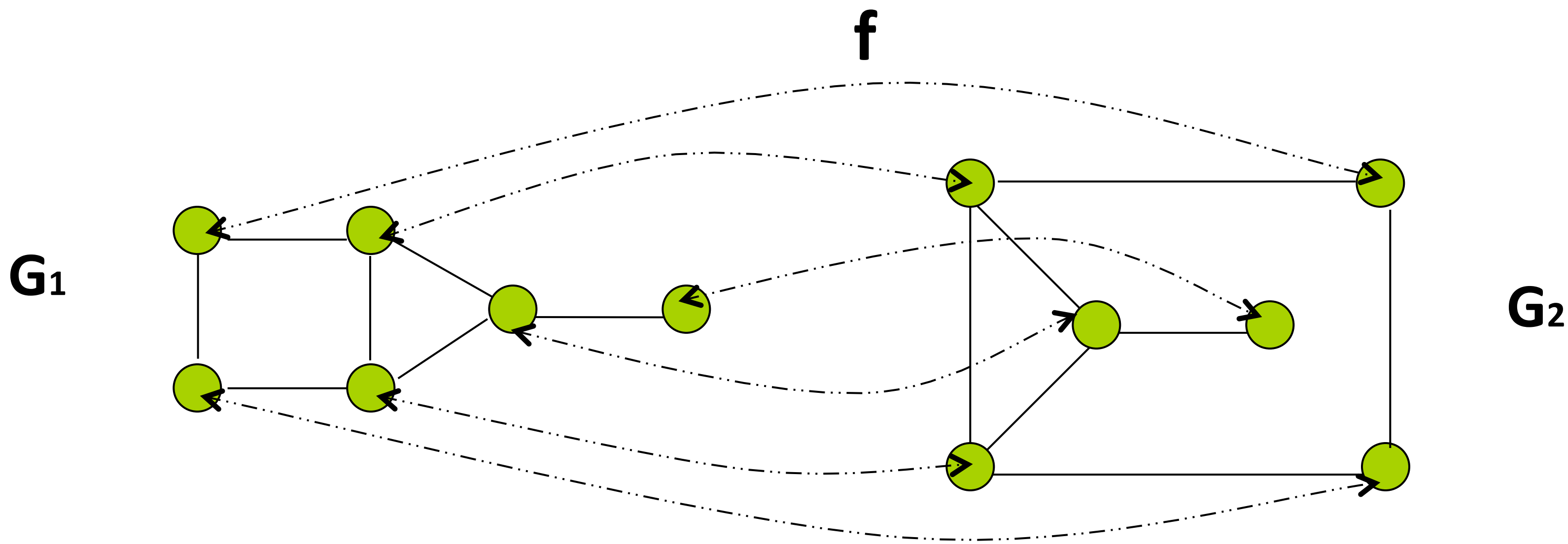


G_2 connesso

Isomorfismo tra grafi

Due grafi G_1 e G_2 sono detti isomorfi se esiste una **relazione biiettiva f: $V_1 \rightarrow V_2$ tra i vertici**.

Ossia c'è una corrispondenza uno a uno tra ogni arco (u_1, v_1) di G_1 e ogni arco (u_2, v_2) di G_2.

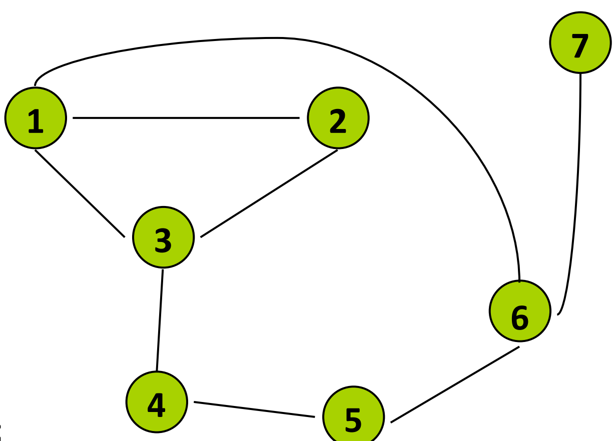


Sottografi

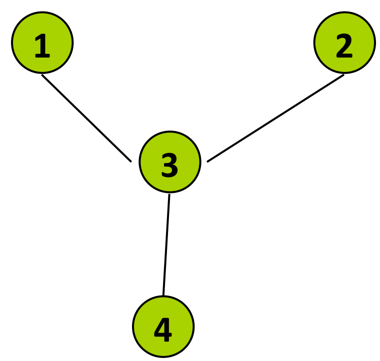
Un grafo G'' = (V'', E'') è un **sottografo** di un grafo G = (V, E) (diretto o non diretto), se si ha V'' $\subseteq$ V e E'' $\subseteq$ E.

Un sottografo di G viene detto **indotto** da V'' quando l'insieme degli archi è definita da tutti che archi di G aventi come vertici incidenti solo quelli appartenenti al sottoinsieme V''. Ossia:

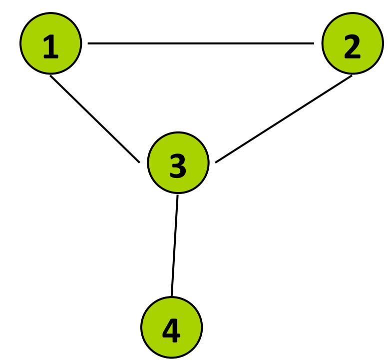
E'' = {(u, v) in E : u, v in V''}



G_1



G_2 sottografo di G_1



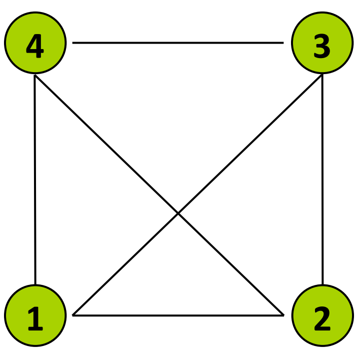
G_3 sottografo indotto da V'' ={1, 2, 3, 4}

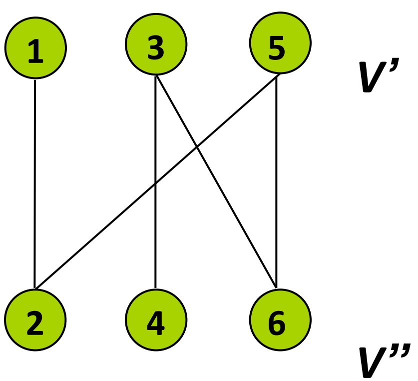
Tipi di grafi

Sia ha un **grafo completo** quando ogni coppia di vertici ha un arco che li unisce.

Un grafo è un **bipartito** quando l' insieme dei vertici V può essere partizionato in due insieme V' e V' tali per cui ogni arco di G è formato da un vertice di V' e uno di V''.

$$(u,v) \in E \quad \text{tale} \quad \text{che} \quad u \in V', v \in V''$$

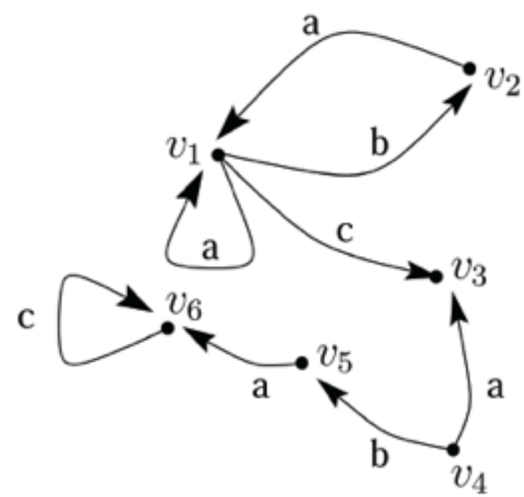




PROPRIETA' DEI GRAFI

Sia G una relazione binaria su un insieme V

1. Se G è *riflessiva* allora il corrispondente grafo avrà un cappio intorno a ogni nodo.
2. Se G è una relazione *irriflessiva* allora nel grafo non ci sono cappi.
3. Se G è una relazione *simmetrica* allora nel grafo, ogni volta che c'è un arco tra due nodi, c'è anche quello che va in direzione opposta.
4. Se G è una relazione *asimmetrica* allora tra due nodi non ci saranno mai un arco e il suo inverso.
5. Se G è una relazione *transitiva* allora nel grafo ogni volta che ci sono due archi "in fila" tra tre nodi x_1, x_2 e x_3 c'è anche l'arco che chiude il triangolo tra x_1 e x_3.



RELAZIONI TRANSITIVE SU GRAFI

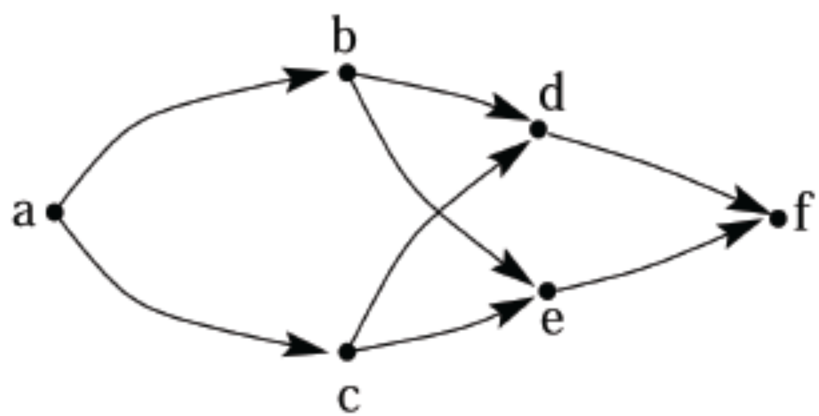
Proposizione 6. *Sia V un insieme e R una relazione binaria su V, $R \subseteq V \times V$. R è transitiva sse ogni qualvolta c'è un cammino di lunghezza $m > 1$ tra due nodi v_i e v_n, allora c'è un arco che collega direttamente v_i a v_n.*

GRAFO DIRETTO ACICLICO (DAG)

Un *grafo diretto aciclico* (detto anche DAG, dall'inglese "Directed Acyclic Graph") è un grafo diretto senza cicli.

Sia $R \subseteq S \times S$ dove $S = \{a, b, c, d, e, f\}$;

$$R = \{\langle a, b\rangle, \langle a, c\rangle, \langle b, d\rangle, \langle c, d\rangle, \langle b, e\rangle, \langle c, e\rangle, \langle d, f\rangle, \langle e, f\rangle, \}.$$



STRUTTURE RELAZIONALI, GRAFI E ORDINAMENTI

(parte 3)

END