INFORMATICA PER LA COMUNICAZIONE

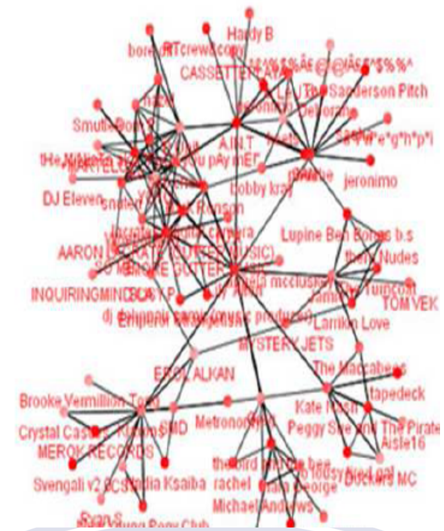
Reti e grafi

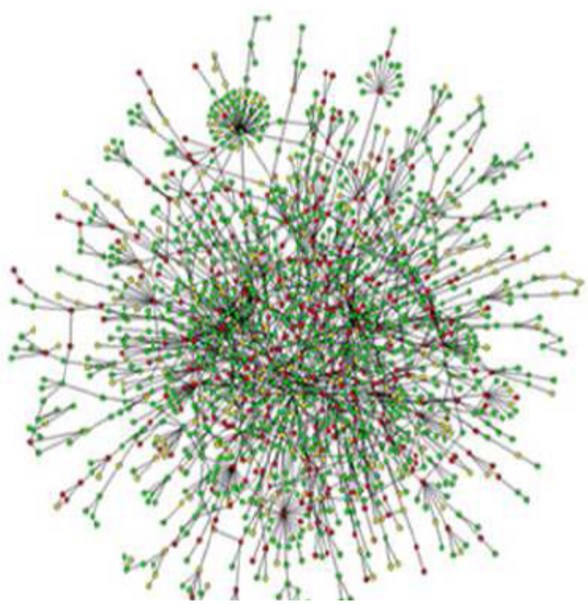
Reti e grafi

- **Reti**: rappresentazione di elementi e delle **possibili relazioni** tra elementi considerati
  - Relazioni **fisiche**
  - Relazioni **immateriali**
- **Internet**: rete **fisica**
- **World Wide Web**: rete **immateriale**

Reti e grafi

- Grafi – Reti: **applicazioni** in diversi ambiti
  - Reti di telecomunicazioni/trasporto
  - Alberi evolutivi
  - Sistemi biologici
  - Reti sociali
  - Rete neurale
  - …





Scienza delle reti – Brevi cenni storici

- Nel 1741 Eulero
  - Risoluzione del problema dei ponti di *Königsberg*
  - Utilizzo tecniche di **teoria dei grafi**
- Nel 1959: modello **grafi casuali** (*Erdos- Renyi*)
- Nel 1999: reti a **invarianza di scala/legge di potenza** (*Barabasi*)

Scienza delle reti – Brevi cenni storici

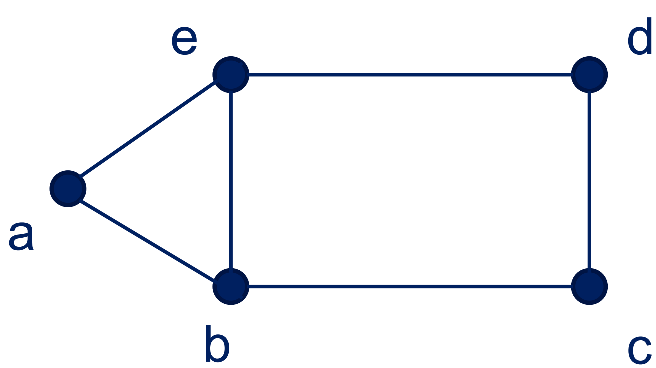
- A partire dalla fine degli anni ‘90:
  - Strumenti per costruire una **mappa** delle reti
  - Dati per l’analisi della struttura delle reti
    - Internet
    - Web
    - Piattaforme social network
    - Reti di proteine
    - …

Teoria dei grafi

Definizioni e applicazioni

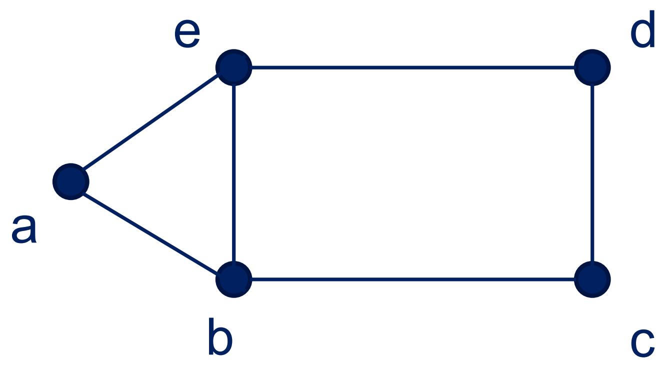
Teoria dei grafi

- **Grafo**: costituito da due tipi di elementi
  - **Nodi** (punti): *a, b, c, d, e*
  - **Archi** collegamento tra coppie di nodi: *(a,e), (a,b), (b,e), (b,c), (d,e), (c,d)*
  - Due nodi *a,b* sono **adiacenti** se esiste l'arco *(a,b)*



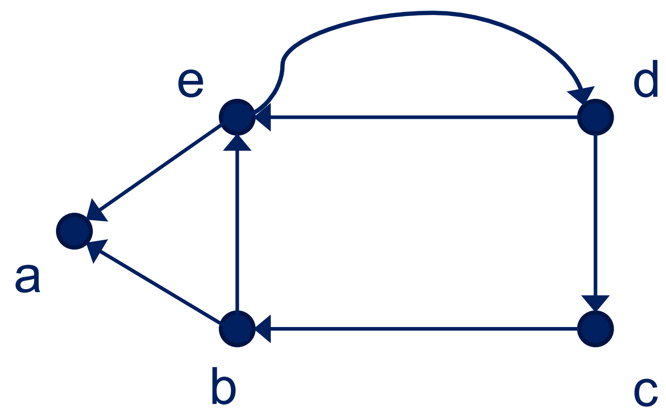
Grafi non diretti

In un grafo **non diretto** ogni arco è **biunivoco**



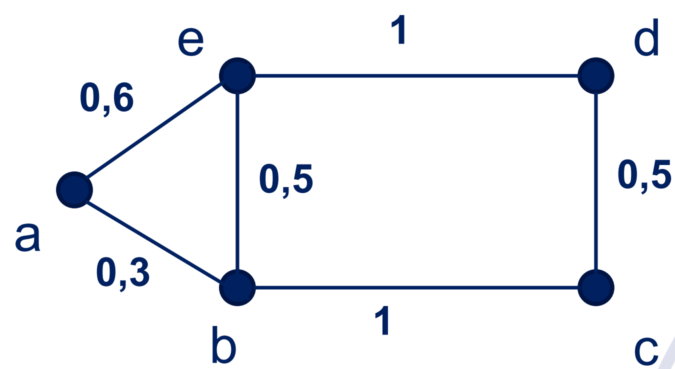
Grafi diretti

- In un grafo **diretto** ad ogni arco viene associata una **direzione**



Grafi pesati

- **Peso** associato a un arco: **intensità** della relazione tra i nodi collegati
- Grafi con archi pesati → **grafi pesati**
- In alcuni casi: pesi associati ai **nodi**

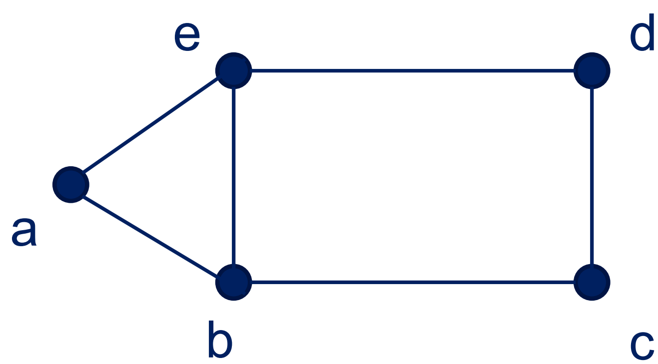


Teoria dei grafi

Archi **incidenti** nel nodo *x*: uno degli estremi in *x*

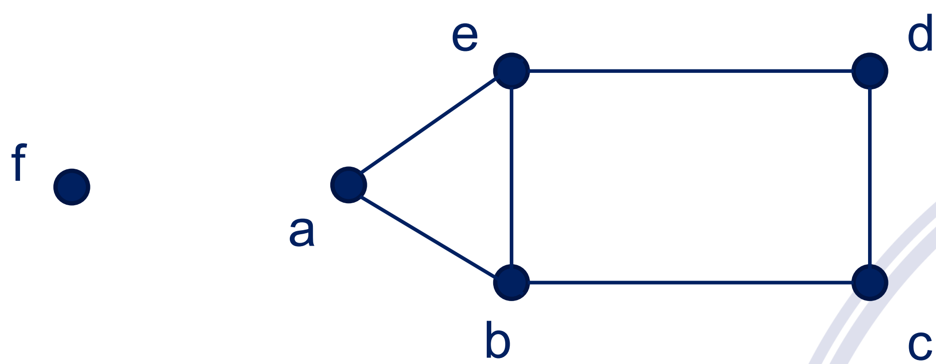
- Archi incidenti in a: *(a,b), (a,e)*
- Archi incidenti in b:

  (a,b), (b,e), (b,c)



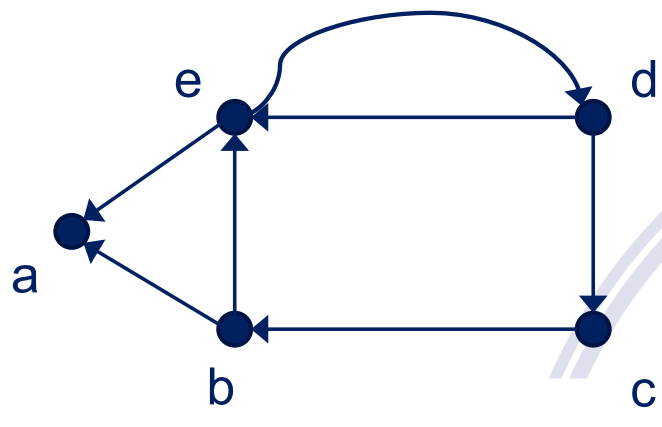
Grado di un nodo

- **Grado** di un nodo (in un grafo *non diretto*): numero di archi incidenti in quel nodo
  - Grado di a = 2
  - Grado di e?
  - Grado di f?



Grado di un nodo

- In un **grafo diretto**
  - Numero di archi diretti verso il nodo (***indegree***)
  - Numero di archi che escono dal nodo (***outdegree***)
    - Nodo *b*: *indegree* = 1, *outdegree* = 2
    - Nodo *a*?
    - Nodo *e*?



Grado medio: rete non diretta

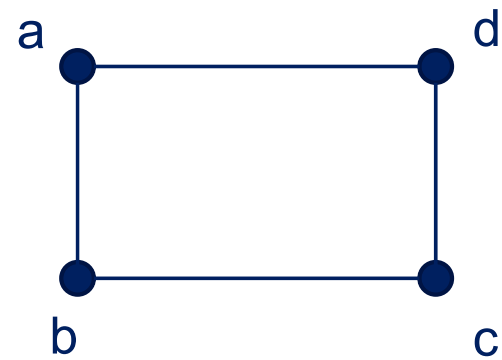
- In un grafo **non diretto**:

$$2E = \sum_{i=1} \mathrm{k}_i \qquad E = \frac{1}{2} \sum_{i=1} \mathrm{k}_i$$

- Il **grado medio** di una rete o di un grafo non diretto

$$\bar{k} = \frac{1}{\mathrm{N}} \sum_{i=1} \mathrm{k}_i = \frac{2E}{N}$$

Grado medio: rete non diretta



- $E = \frac{1}{2}\sum_{i=1} \mathrm{k}_i = \frac{1}{2}\,8 = 4$

- Il **grado medio** di una rete o di un grafo non diretto

$$\bar{k} = \frac{1}{\mathrm{N}}\sum_{i=1} \mathrm{k}_i = \frac{2E}{N} = \frac{8}{4} = 2$$

Grado medio: rete diretta

- In un grafo **diretto**:

$$E = \sum_{i=1} \mathrm{k}^{+}{}_{i} = \sum_{i=1} \mathrm{k}^{-}{}_{i}$$

- Il **grado medio** di una rete o di un grafo diretto

$$\bar{k} = \frac{1}{\mathrm{N}} \sum_{i=1} \mathrm{k}^{+}{}_{i} = \frac{E}{N}$$

Reti esistenti

NETWORK	NODES	LINKS	DIRECTED UNDIRECTED	N	E	$\bar{k}$
Internet	Routers	Internet connections	Undirected	192,244	609,066	6.34
WWW	Webpages	Links	Directed	325,729	1,497,134	4.60
Power Grid	Power plants, transformers	Cables	Undirected	4,941	6,594	2.67
Mobile Phone Calls	Subscribers	Calls	Directed	36,595	91,826	2.51
Email	Email addresses	Emails	Directed	57,194	103,731	1.81
Science Collaboration	Scientists	Co-authorship	Undirected	23,133	93,439	8.08
Actor Network	Actors	Co-acting	Undirected	702,388	29,397,908	83.71
Citation Network	Paper	Citations	Directed	449,673	4,689,479	10.43
E. Coli Metabolism	Metabolites	Chemical reactions	Directed	1,039	5,802	5.58
Protein Interactions	Proteins	Binding interactions	Undirected	2,018	2,930	2.90

Da A. L. Barabasi, *Network Science*

Distribuzione del grado

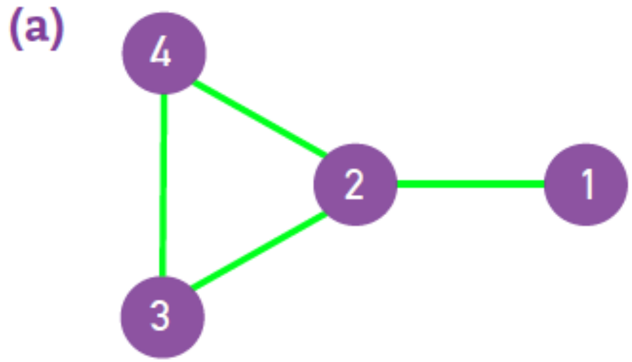
- **Distribuzione del grado**: **porzione** di nodi con grado *k*
- Posso considerarla come una **distribuzione di probabilità**
- Essendo una probabilità:

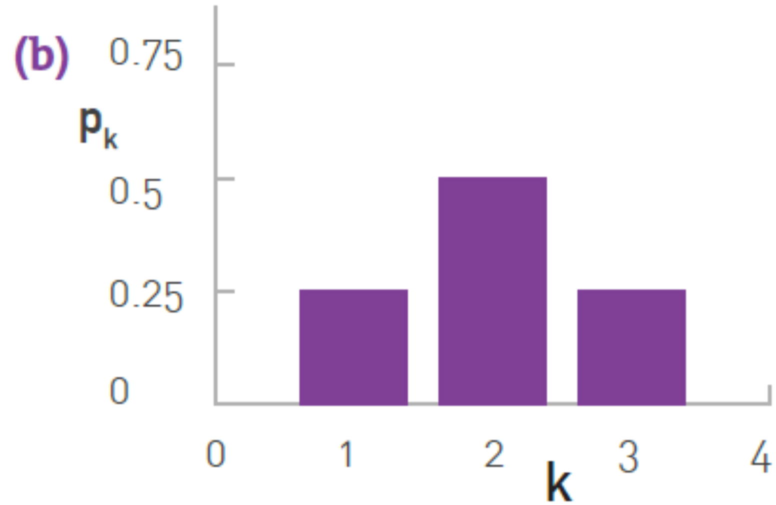
$$\sum_{k=0} p_k = 1$$

- La probabilità che un nodo abbia grado *k* è:

$$p_k = \frac{N_k}{N} \quad \text{e } N_k = p_k N$$

Distribuzione del grado





Da A. L. Barabasi, *Network Science*

Grado medio e distribuzione

- Considerando il grado dei nodi e la distribuzione del grado, posso definire il **grado medio** di una rete:

$$\bar{k} = \sum_{i=0} i \, p_i$$

Grado medio e distribuzione

Esercizio: Abbiamo un grafo con 5 nodi dove ogni nodo ha un grado che varia da 1 a 3. Supponiamo che i nodi siano distribuiti in questo modo:

- 1 nodo ha grado 1
- 3 nodi hanno grado 2
- 1 nodo ha grado 3

- **Il grado medio?**

Grado medio e distribuzione

Soluzione: La probabilità p_i di ogni grado si calcola dividendo il numero di nodi con quel grado per il numero totale di nodi. Quindi avremo:

$p_1 = \frac{1}{5}$ perché c'è 1 nodo con grado 1

$p_2 = \frac{3}{5}$ perché ci sono 3 nodi con grado 2

$p_3 = \frac{1}{5}$ perché c'è 1 nodo con grado 3

Il grado medio $\bar{k}$ si calcola moltiplicando il grado per la probabilità corrispondente e sommando i risultati:

$$\bar{k} = \left(1 \, . \, \frac{1}{5}\right) + \left(2 \, . \, \frac{3}{5}\right) + \left(3 \, . \, \frac{1}{5}\right) = \frac{1}{5} + \frac{6}{5} + \frac{3}{5} = 2$$

Grado medio e distribuzione

Esercizio: Abbiamo un grafo con 6 nodi dove ogni nodo ha un grado che varia da 1 a 4. Supponiamo che i nodi siano distribuiti in questo modo:

- 2 nodi hanno grado 1
- 1 nodo ha grado 2
- 2 nodi hanno grado 3
- 1 nodo ha grado 4

- **Il grado medio?**

Grado medio e distribuzione

Soluzione: La probabilità p_i di ogni grado si calcola dividendo il numero di nodi con quel grado per il numero totale di nodi. Quindi avremo:

$p_1 = \frac{2}{6}$ perché ci sono 2 nodi con grado 1

$p_2 = \frac{1}{6}$ perché c'è 1 nodo con grado 2

$p_3 = \frac{2}{6}$ perché ci sono 2 nodi con grado 3

$p_4 = \frac{1}{6}$ perché c'è 1 nodo con grado 4

Il grado medio $\bar{k}$ si calcola moltiplicando il grado per la probabilità corrispondente e sommando i risultati:

$$\bar{k} = \left(1 \,.\, \frac{2}{6}\right) + \left(2 \,.\, \frac{1}{6}\right) + \left(3 \,.\, \frac{2}{6}\right) + \left(4 \,.\, \frac{1}{6}\right) = \frac{14}{6} = 2.33$$

Matrice di adiacenza

Per rappresentare la **topologia** di un grafo:

- **Matrice di adiacenza**: matrice binaria per rappresentare presenza/assenza archi

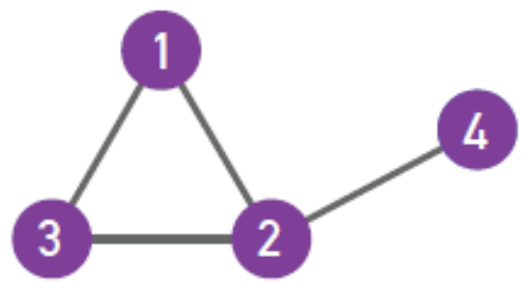
(a) **Adjacency matrix**

$$A_{ij} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix}$$

Da A. L. Barabasi, *Network Science*

Matrice di adiacenza

(b) **Undirected network**



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

Da A. L. Barabasi, *Network Science*

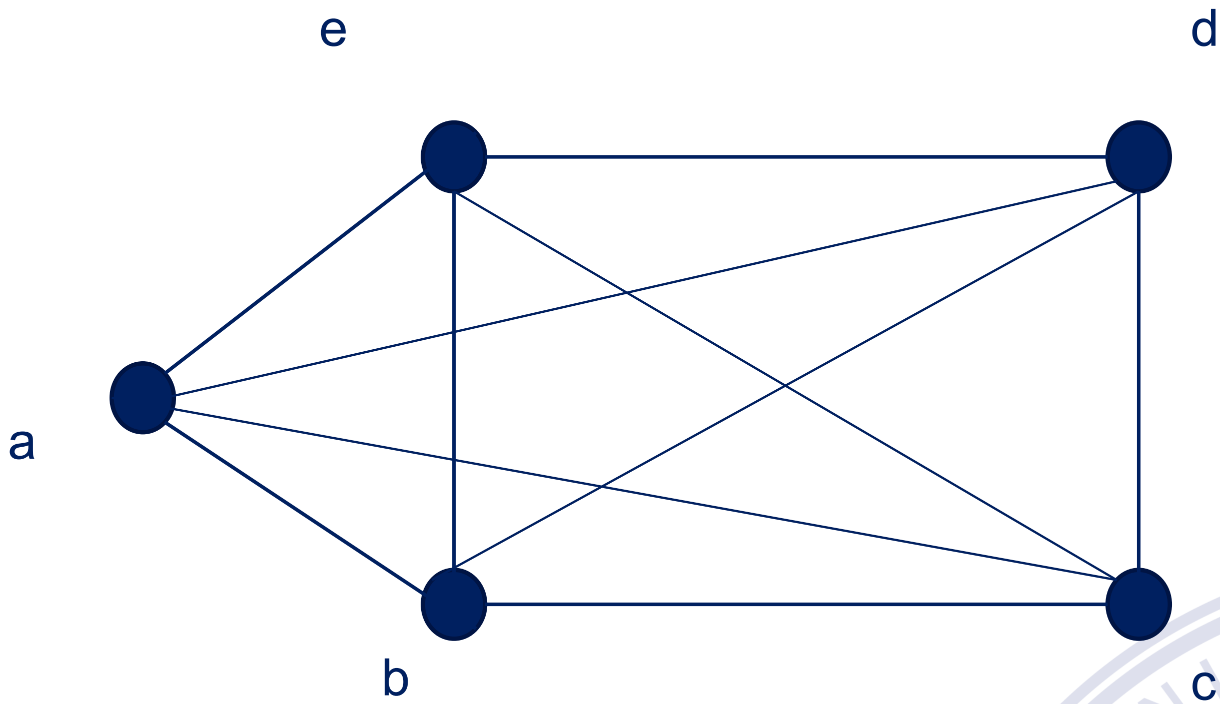
Grado massimo

- Il **grado** di un nodo è **al più *N-1***
- In un grafo **non diretto**:

$$E \leq \frac{\mathrm{N(N-1)}}{2}$$

- Un grafo con *N(N-1)/2* archi è un **grafo completo**

Grafo completo



Grado e reti esistenti

- Se tutti i **link** fossero presenti nella rete WWW: *$5*10^{10}$ archi*
- Nella realtà circa *$15*10^{5}$ archi*
- Esistono circa *$3*10^{-5}$* dei possibili archi

NETWORK	NODES	LINKS	DIRECTED UNDIRECTED	N	E	$\bar{k}$
Internet	Routers	Internet connections	Undirected	192,244	609,066	6.34
WWW	Webpages	Links	Directed	325,729	1,497,134	4.60

Da A. L. Barabasi, *Network Science*

Reti sparse

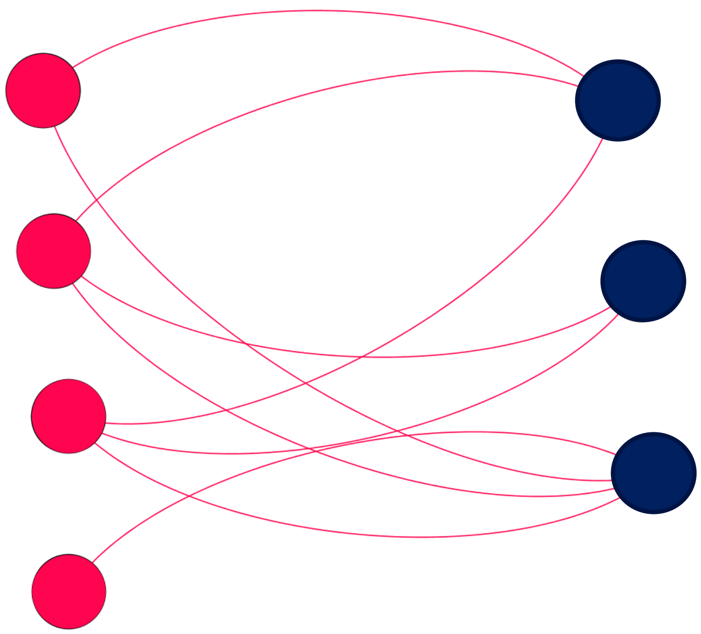
- Molte delle reti reali sono **sparse**: numero **limitato** di archi
- Di conseguenza, anche le **matrici di adiacenza sono sparse**
- Esiste una modalità più efficiente di memorizzazione: **liste di adiacenza**

Grafi bipartiti

Definizione e proprietà

Grafo bipartito

- **Grafo bipartito**: i nodi appartengono a due insiemi/categorie
- **Non** vi sono collegamenti **interni** a un insieme



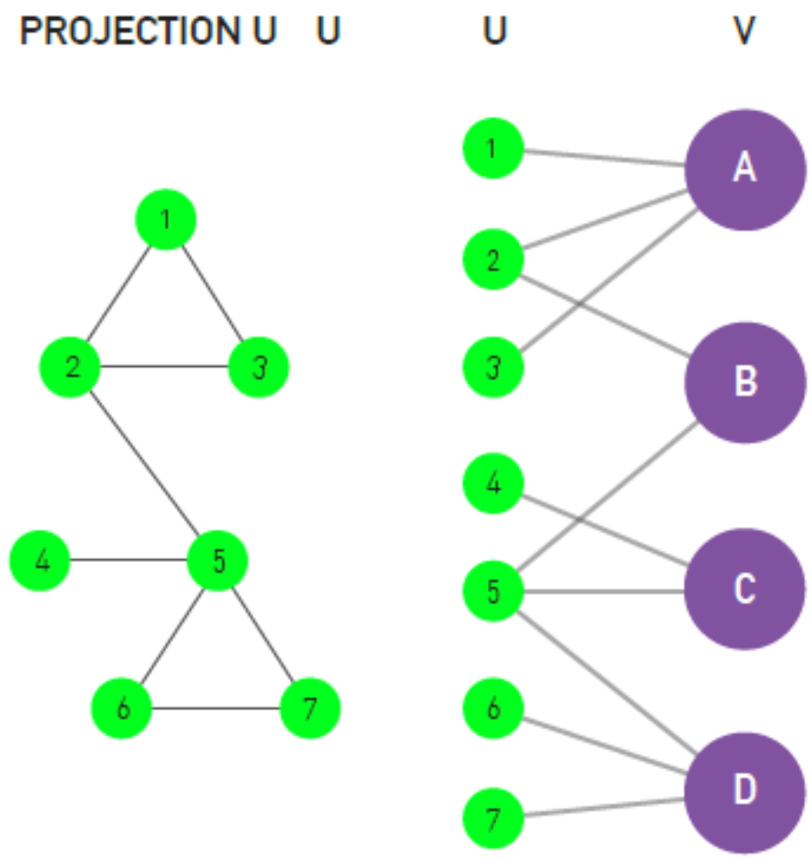
Grafo bipartito

- **Grafo bipartito**: due categorie di elementi
  - Utenti – siti visitati
  - Consumatori – merci acquistate
  - Attori – film
  - ….

Grafo bipartito e proiezioni

- A partire dal grafo bipartito: **proiezioni**
- **Proiezione**: grafo (non necessariamente bipartito)
  - **Nodi**: elementi di un insieme
  - **Archi**: esiste un collegamento se due elementi sono collegati a un elemento comune

Grafo bipartito e proiezioni



Da A. L. Barabasi, *Network Science*

Grafi multipartiti

- Un grafo multipartito è costituito da diversi insiemi (2, 3, ...)
  - Collegamenti tra nodi di **insiemi differenti**
  - **Nessun collegamento interno** agli insiemi

Cammini e distanze

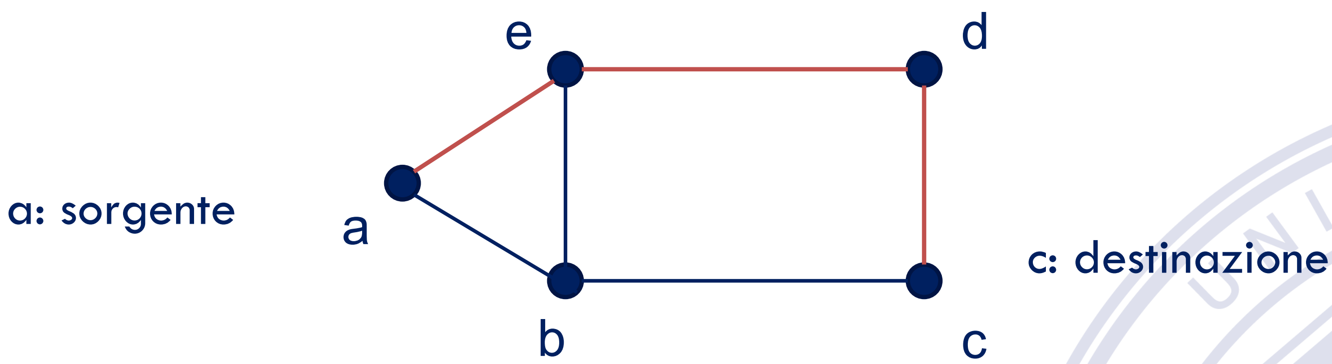
Definizioni e proprietà

Distanze e cammini

- Determinare la **distanza** tra due punti in una rete è di importanza fondamentale:
  - Distanza tra due città
  - Distanza tra due utenti di una piattaforma
  - Distanza tra due pagine web
  - …
- Per definire la distanza in una rete: definizione di **cammino**

Cammino

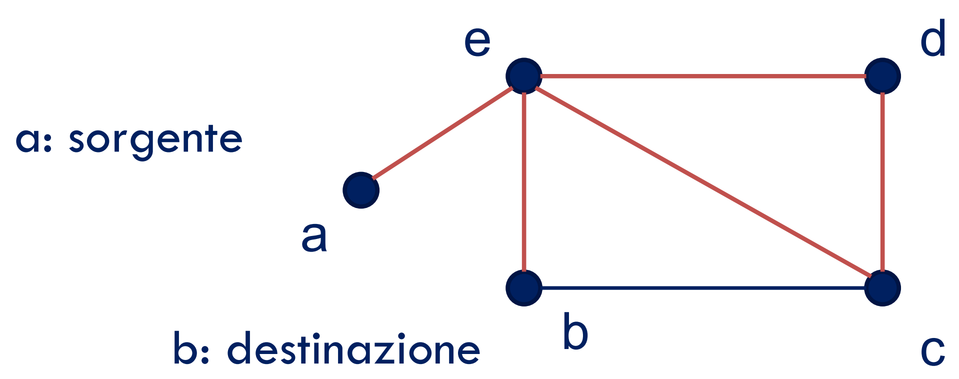
- **Cammino** in un grafo: **sequenza di nodi**
  - **Adiacenti**
  - **Distinti**
  - Nodo **sorgente**
  - Nodo **destinazione**



Cammino *a*, *e*, *d*, *c*

Percorso

- **Percorso** in un grafo: **sequenza di nodi**
  - **Adiacenti**
  - Nodo **sorgente**
  - Nodo **destinazione**
- Nodi possono essere **ripetuti**

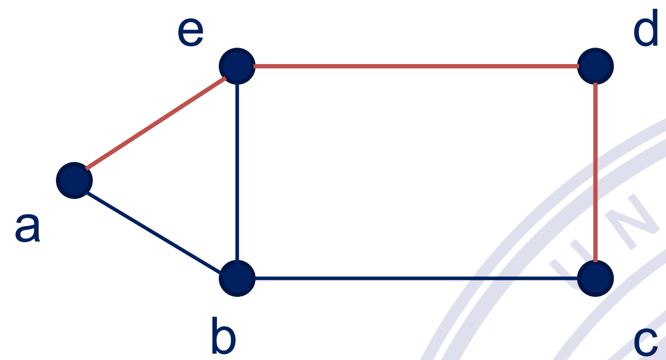


Percorso a, **e**, d, c, **e**, b

Cammini e percorsi

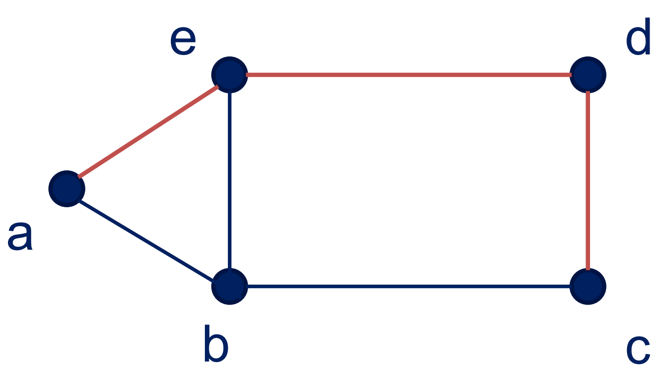
- **Lunghezza** di un cammino (percorso): *numero di archi* tra sorgente e destinazione
- Cammino *a, e, d, c*: lunghezza 3

Cammino *a, e, d, c*

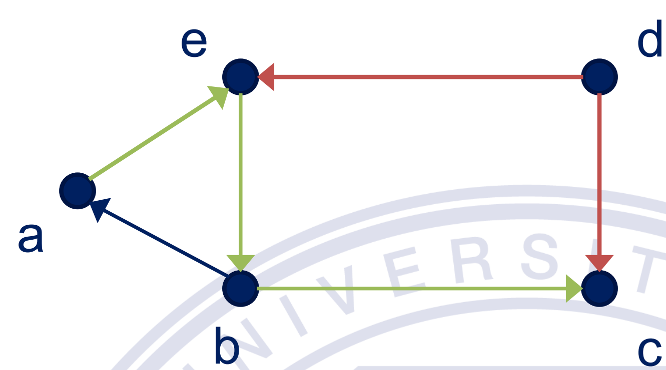


Cammini e percorsi

- In un grafo **diretto** un percorso deve seguire la **direzione** degli archi
  - Differenza tra grafi diretti e non diretti



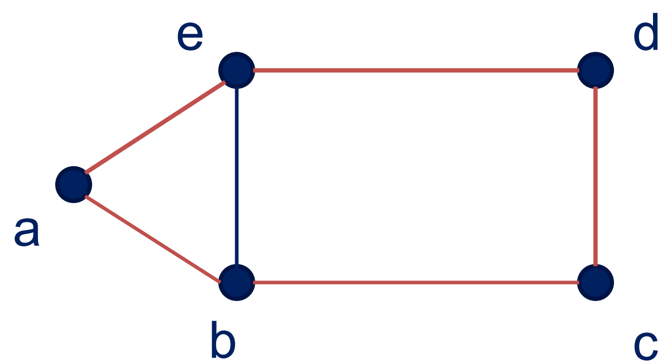
Cammino *a, e, d, c*



Cammino *a, e, b, c*

Ciclo

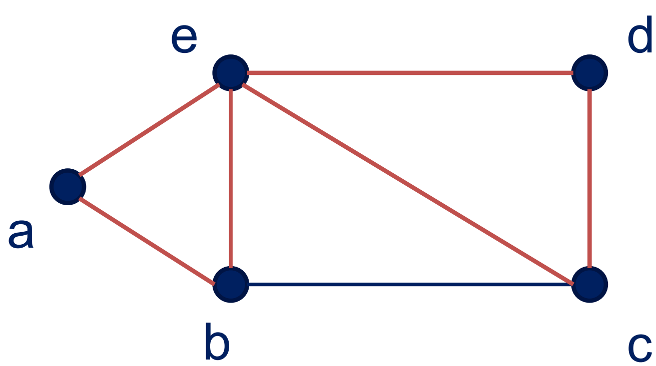
- **Ciclo: percorso** i cui nodi sono **distinti eccetto il primo e l'ultimo**



Ciclo *a*, e, *d*, *c*, *b*, *a*

Circuito

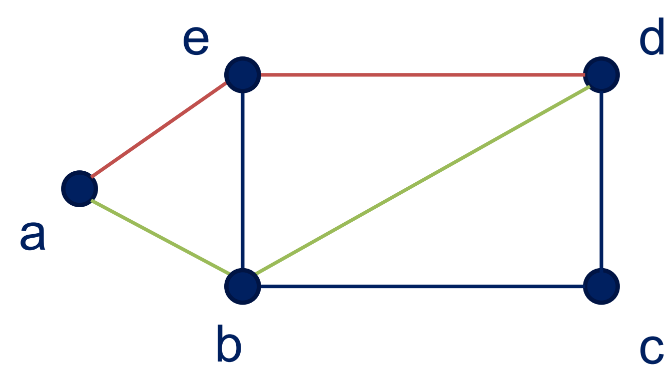
- **Circuito**: percorso in cui **il primo e l'ultimo nodo coincidono**



Circuito *a*, ***e***, *d*, *c*, ***e***, *b*, *a*

Cammino minimo

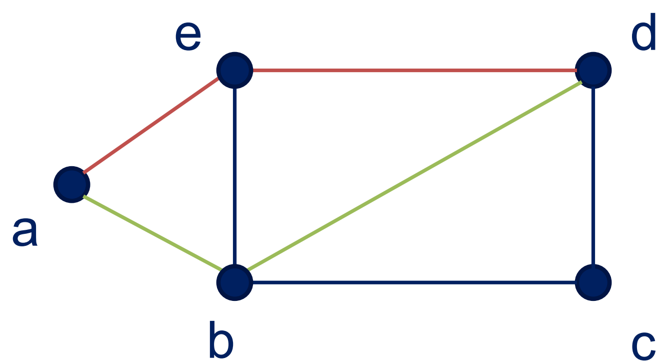
- Un **cammino** tra due nodi è **minimo** quando *nessun altro cammino tra i due nodi è composto da un numero inferiore di archi*
- Possono esistere più cammini minimi



Cammini minimi tra *a* e *d*: (1) *a, e, d* (2) *a, b, d*

Distanza

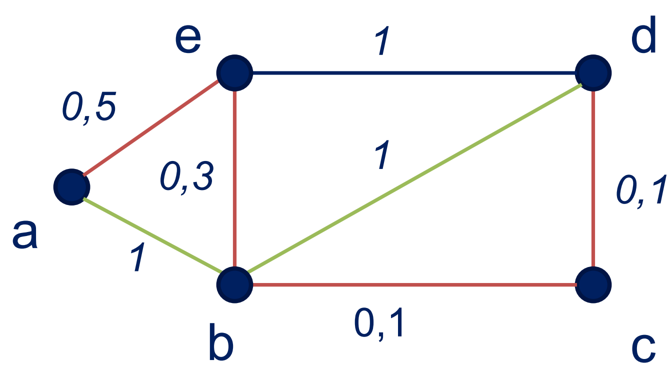
La **distanza** tra due nodi *u* e *v* è la lunghezza del **cammino minimo** tra *u* e *v*



Distanza tra *a* e *d*: 2

Cammino

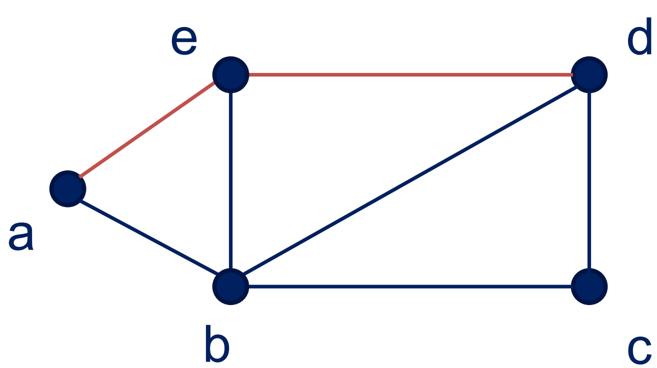
- **Cammino** (minimo): caso **pesato** → somma dei **pesi** sugli archi



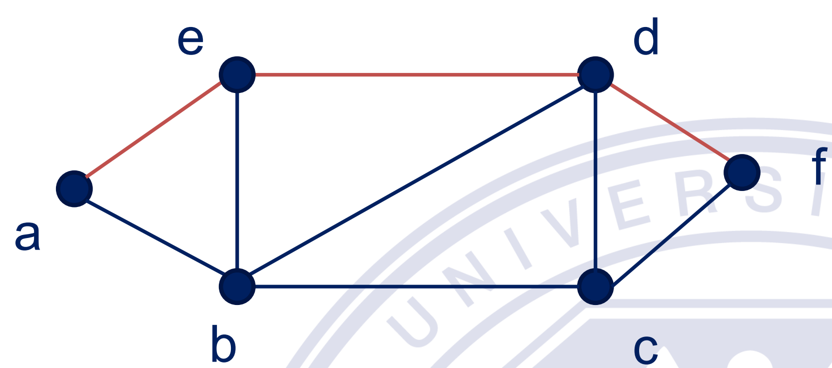
Cammino minimo tra *a* e *d*: *a, e, b, c, d* costo: *1*

Diametro

- Il **diametro** di un grafo è costituito dal valore della **massima distanza**, cioè dal numero di archi che compongono il *cammino minimo più lungo* presente nel grafo



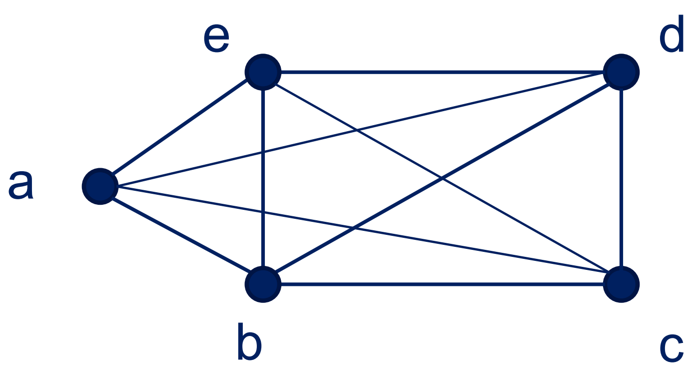
Diametro = 2



Diametro = 3

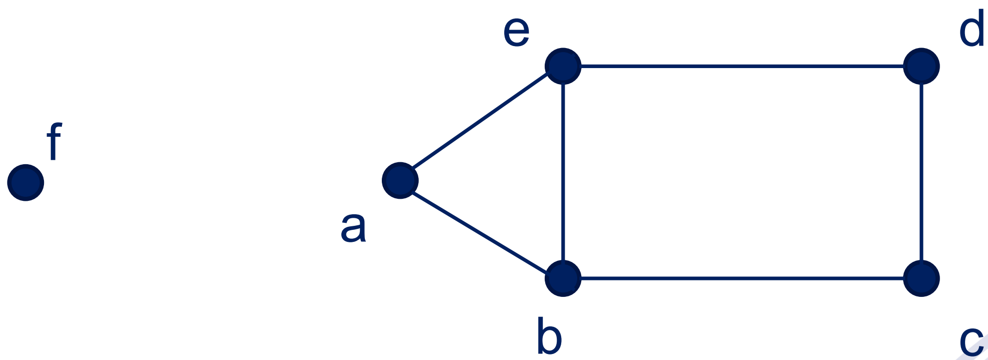
Teoria dei grafi

Diametro?



Teoria dei grafi

Diametro?



Distanza media

- In un grafo possiamo considerare anche la **distanza media**:

$$\overline{dist} = \frac{1}{\mathrm{N}(\mathrm{N}-1)} \sum_{i,\mathrm{j}} \mathrm{d}ist_{i,j}$$

Distanza media

Esercizio: Immagina un grafo con 5 nodi etichettati A, B, C, D ed E, collegati come segue:

- A è connesso a B e D
- B è connesso a A, C ed E
- C è connesso a B e D
- D è connesso a A, C ed E
- E è connesso a B e D

Distanza media ?

Distanza media

Soluzione: Le distanze tra ciascuna coppia sono:

A a B = 1	A a C = 2	A a D = 1	A a E = 2
B a C = 1	B a D = 2	B a E = 1	
C a D = 1	C a E = 2		
D a E = 1			

distanza media = 14/20 = 0.7

Connettività

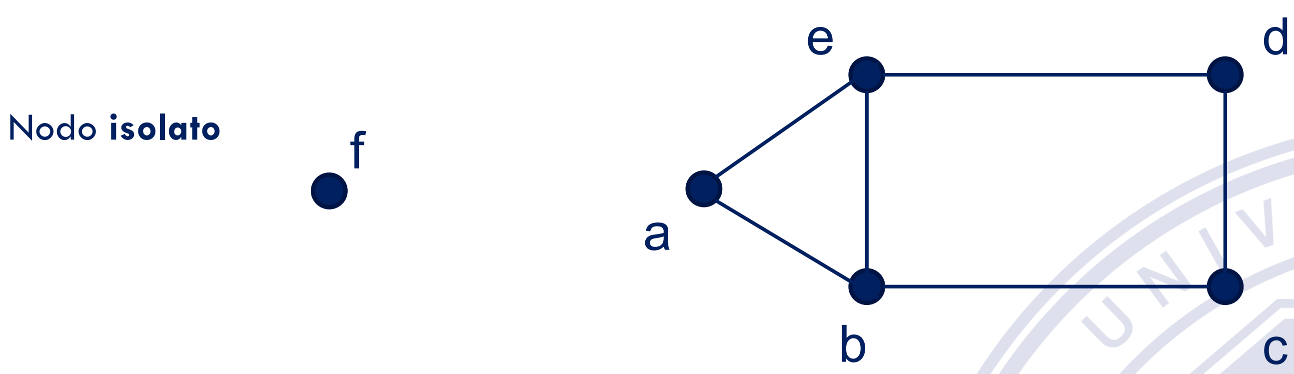
Definizioni e proprietà

Connessioni

- Una delle proprietà fondamentali delle reti: **possibilità di comunicare** tra elementi del sistema
  - Connessione tra elaboratori
  - Connessione tra due utenti
  - …
- Quali elementi possono essere messi in comunicazione?

Componente connessa

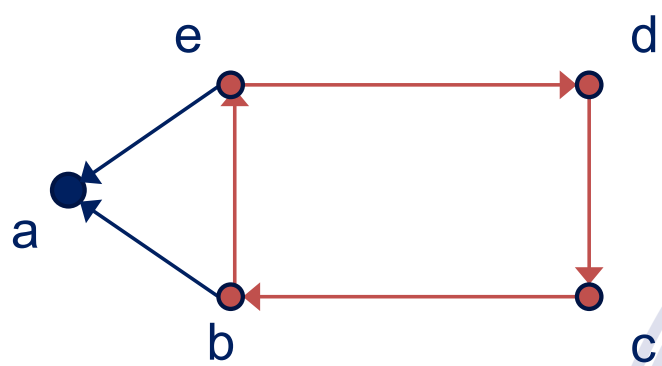
- **Componente connessa**: insieme dei nodi del grafo
  - Per ogni coppia di nodi esiste un cammino che li collega
- Grafo **connesso**: esiste una sola componente connessa
- Grafi diretti/ Grafi non diretti?



Componenti connesse: (1) *f* (2) *a, b, c, d, e*

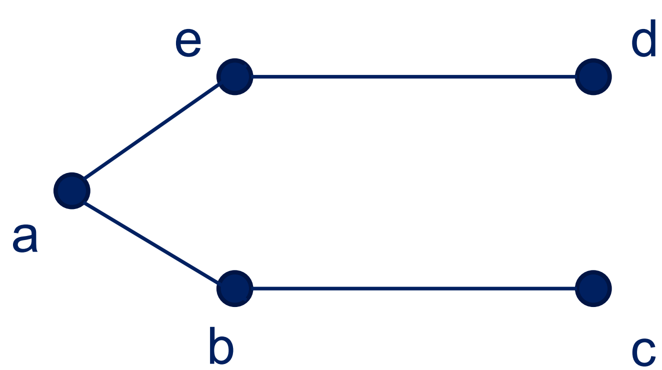
Componente fortemente connessa

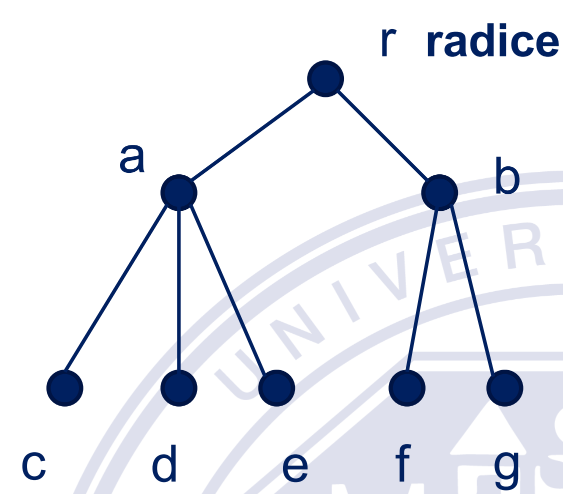
- Nei grafi diretti si parla di componente **fortemente connessa**: da ogni nodo della componente è possibile raggiungere gli altri nodi della componente



Alberi

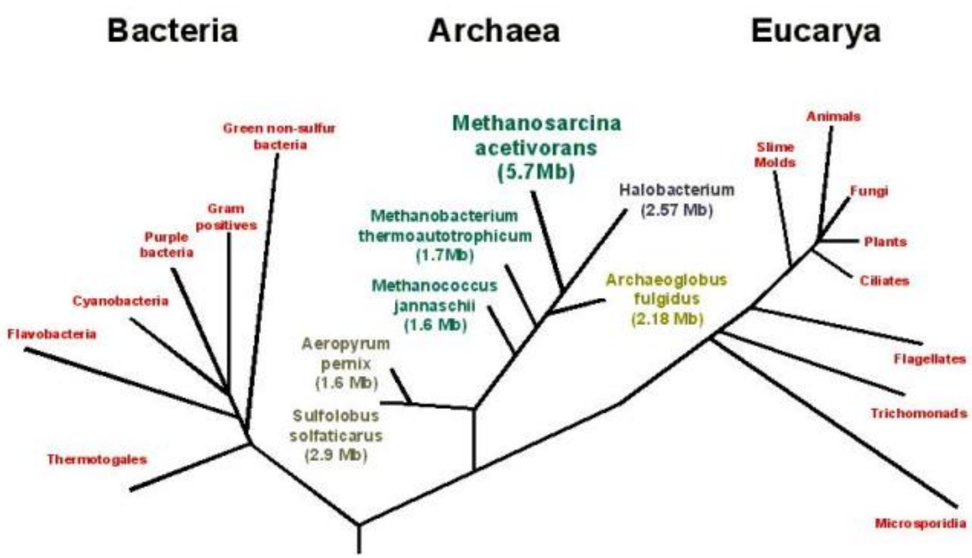
- **Albero**: grafo connesso **privo di cicli**
  - Tra ogni coppia di nodi esiste **uno e un solo cammino**
  - **Foglie**: nodi di grado 1
  - Presenza o meno di **radice**





Alberi

- Applicazioni **alberi**:
  - Alberi evolutivi
  - **Procedure decisionali**

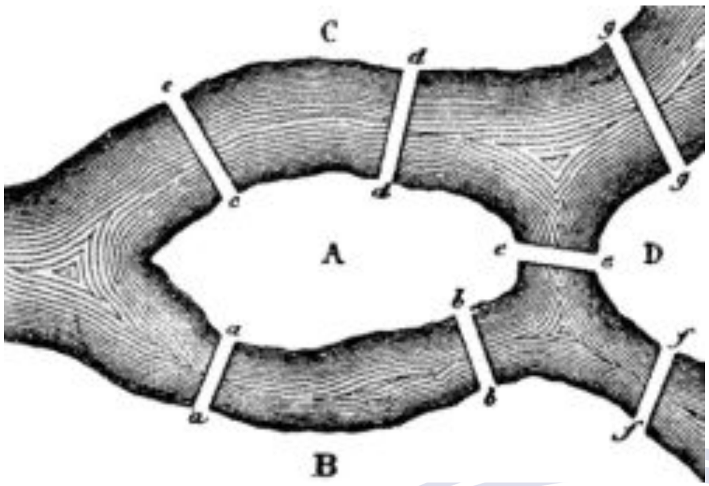


Ponti di Könisberg

Problema e soluzione

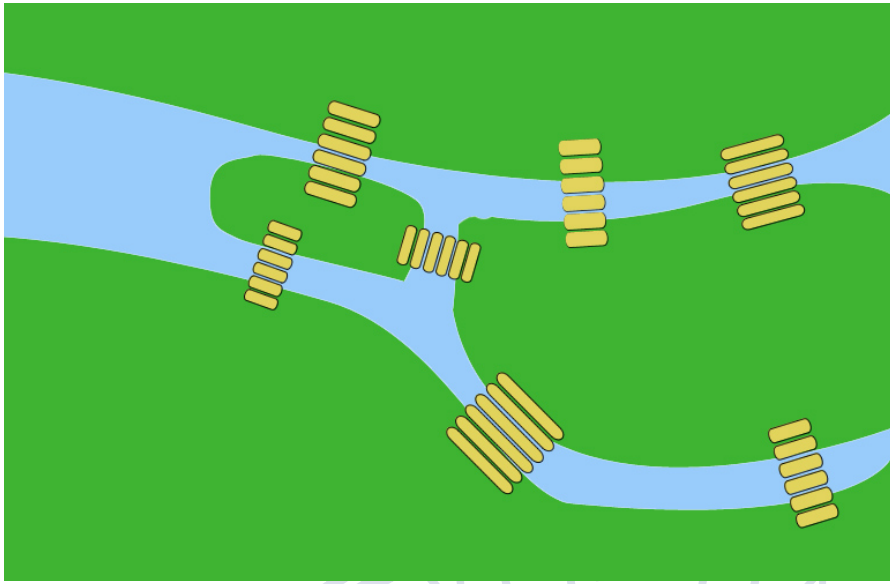
Ponti di Königsberg

- Königsberg (Kalingrad): città natale di Kant sul fiume Pregel
- **Sette** ponti attraversano il fiume Pregel



Ponti di Könisberg

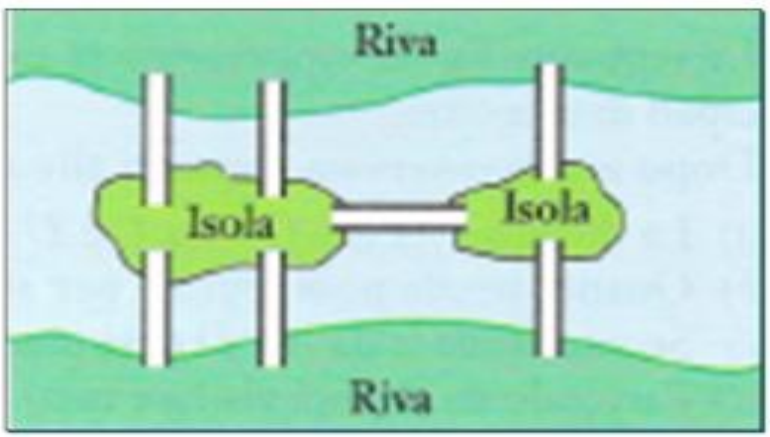
Problema: esiste un modo per attraversare tutti i ponti percorrendone ognuno una e una sola volta?



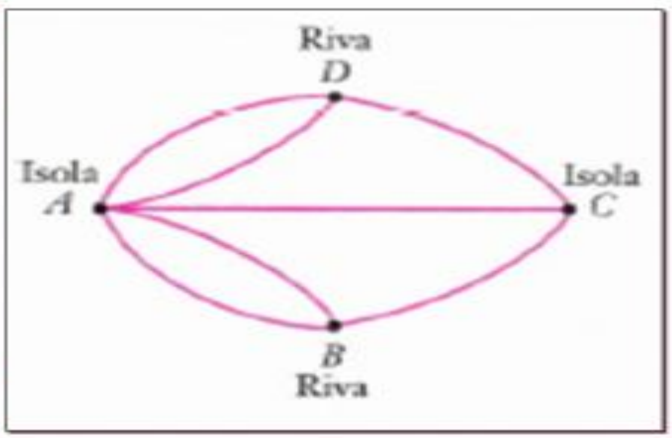
Ponti di Könisberg

Rappresentazione del problema tramite un **grafo**:

- **Nodi**: **parti** della città
- **Archi**: **ponti**



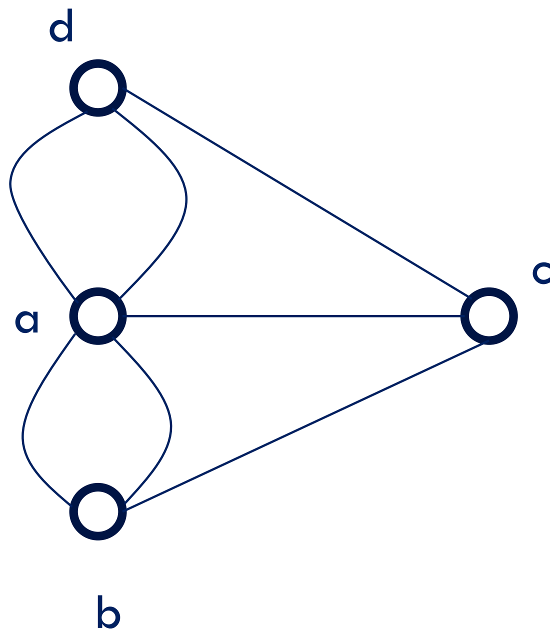
I sette ponti di Königsberg



Schematizzazione mediante una rete topologica

Ponti di Könisberg

Il problema ammette una soluzione?



Ponti di Könisberg

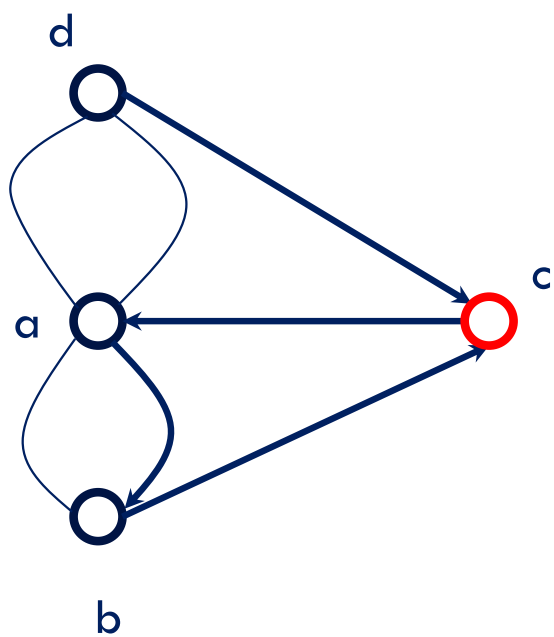
Impossibilità attraversamento sette ponti (Eulero, 1741)

Consideriamo un percorso che attraversa i ponti

1. Esiste un **nodo intermedio** nel percorso
2. Consideriamo gli archi che attraversano quel nodo
   - **Tutti** gli archi sono percorsi
   - Complessivamente sono percorsi un **numero pari** di volte
3. Tutti i nodi hanno grado **dispari**

Ponti di Könisberg

Consideriamo il nodo *c*



Teoria dei grafi

Proprietà generale: *esiste un percorso che attraversa tutti gli archi di un grafo una e una sola volta, solo se il grafo contiene **al più due nodi di grado dispari***.

Ogni nodo intermedio deve essere attraversato un numero pari di volte

Ponti di Könisberg

Aggiunta di un arco