



---

ALMA MATER STUDIORUM A.D. 1088  
**UNIVERSITÀ DI BOLOGNA**

AA 2013-2014 - Corso 44756 - 6 CFU

# ANALISI DELLE RETI SOCIALI

## *Social Network Analysis*

### **PARTE I – FONDAMENTI DI MISURE DI RETE**

#### **DISPENSA DI SUPPORTO ALLA DOCENZA**

Rev. al 26 settembre 2013

marco ruffino [marco.ruffino@unibo.it](mailto:marco.ruffino@unibo.it)

## SOMMARIO

<b>Come nasce e a cosa serve la <i>Social Network Analysis</i></b> .....	<b>3</b>
<b>Tipologie di rete ed operazioni sui dati</b> .....	<b>8</b>
Tipologie di reti .....	8
Scale di misura .....	10
Modalità di rappresentazione dei dati .....	11
Operazioni sulle matrici .....	15
<b>Misure di rete</b> .....	<b>20</b>
Misure generali e di connessione .....	20
Misure di centralità .....	27
Misure relative ad Ego .....	32
Misure relative ai gruppi .....	35
Misure relative a ruoli e posizioni .....	38
<b>Bibliografia essenziale</b> .....	<b>47</b>

## COME NASCE E A COSA SERVE LA SNA /1

### I PRINCIPI

*“L'uomo è un essere sociale interagente, capace di manipolare gli altri così come di essere manipolato da loro. [...] Il postulato fondamentale del network approach è che le persone sono viste come interagenti con altre persone, alcune delle quali a loro volta interagiscono fra di loro e con altri ancora, e che il network totale di relazioni che così si forma è in uno stato di fluidità”* (Boissevain J., *Network analysis. Studies in Human Interaction*, Paris, 1973).

- *Network approach*: alcuni principi “ideologici” ...
  - Le relazioni sociali strutturate sono un mezzo più potente di spiegazione sociologica di quanto non lo siano gli attributi personali dei membri del sistema.
  - Le norme derivano dalla posizione nei sistemi strutturati di relazioni sociali
  - Le strutture sociali determinano l'attività delle relazioni diadiche
  - il mondo è composto di networks, non di gruppi
  - I metodi strutturali integrano e sostituiscono i metodi individualistici
- ... ed analitici:
  - i legami di solito sono reciproci in maniera asimmetrica, perché differiscono nel contenuto e nell'intensità;
  - i legami uniscono i membri del network indirettamente e direttamente. Di conseguenza, essi devono essere definiti entro il contesto di strutture di network più larghe;
  - lo strutturarsi dei legami sociali crea networks non casuali: clusters, confini e legami incrociati;
  - i legami trasversali uniscono i clusters così come gli individui;
  - legami asimmetrici e networks complessi distribuiscono in maniera differenziata risorse scarse.

## COME NASCE E A COSA SERVE LA SNA /2

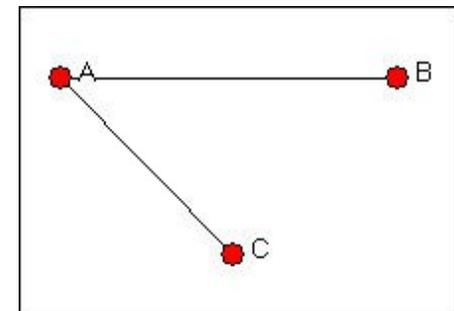
### LE LOGICHE

- **Approccio strutturalista** (ovvero topologico, basato sulle equivalenze di posizione, in una visione della rete come “prisma”). Due nodi hanno gli stessi comportamenti o ottengono gli stessi risultati in quanto occupano nella rete una posizione simile, anche se fra loro non hanno connessioni. L'omogeneità di comportamento è dunque figlia della posizione in cui un soggetto si trova, cioè dell'ambiente sociale. Un approccio strutturalista più complesso considera l'equivalenza strutturale come condizione che permette ad ogni attore di riconoscere gli altri attori a lui simili (anche se non li ha mai conosciuti) e/o di imitarne i comportamenti (isomorfismo mimetico).

Esempio: nell'approccio strutturalista, la rete proposta di lato (“buco strutturale”, vedi oltre) è interpretata come “possibilità di A di giocare B contro C”.

- **Approccio connessionista** (ovvero basato su flussi, relazioni, grado di coesione, in una visione della rete come “tubi”, rivolta alla valutazione dei benefici informativi). Il connessionismo implica un processo di trasmissione interpersonale fra legami sociali preesistenti, utilizzando meccanismi di modellazione e congruenza.

Esempio: nell'approccio connessionista, la rete proposta di lato è interpretata come possibilità di A di massimizzare la quantità di informazione non ridondante che riceve attraverso i suoi contatti derivante in quanto questi non sono connessi fra di loro.



## COME NASCE E A COSA SERVE LA SNA /3

### LE LOGICHE

- Due diverse prospettive per leggere gli stessi oggetti:
  - **Focus sulla varietà:** gli approcci basati sulle performance date dal **capitale sociale**, che consente ai singoli attori di definire strategie e costruire “organizzazioni” in funzione della loro dotazione di legami. I legami cioè “abilitano” gli individui a compiere le loro strategie. Dal punto di vista della ricerca, si cerca di istituire relazioni esplicative fra “successo” (variazione di attributi) e legami di cui l'attore dispone.
  - **Focus sull'omogeneità:** gli approcci basati sugli impatti “normalizzanti” del network sui comportamenti dei singoli attori che ne fanno parte, in funzione del tipo di legami di cui dispongono. Dal punto di vista della ricerca, si pone in correlazione l'omogeneità delle attitudini e dei comportamenti degli attori con la natura dei loro legami.
- Oltre la sociologia, la network analysis in campo economico:
  - la rete come modello dell'organizzazione dell'impresa e dei mercati, rileggendo i concetti di “costo di transazione” (costi di accesso) e “dipendenza dalle risorse” (asimmetria, alleanze), oltre che ovviamente di riduzione dell'incertezza (controllo attraverso interdipendenza);
  - un approccio emergente: i network come **ambienti cognitivi**, in quanto spazio delle interazioni economiche e sociali attraverso cui scorrono significati, varietà e conoscenza. Alcuni riferimenti sono al *learning by interacting* (accesso ad informazioni e conoscenza altrimenti non disponibile), alle *cognitive maps* individuali e collettive, alla “memoria transattiva”, in genere al *knowledge management*.

## COME NASCE E A COSA SERVE LA SNA /4

### LA CRESCITA DELLA SNA IN LETTERATURA

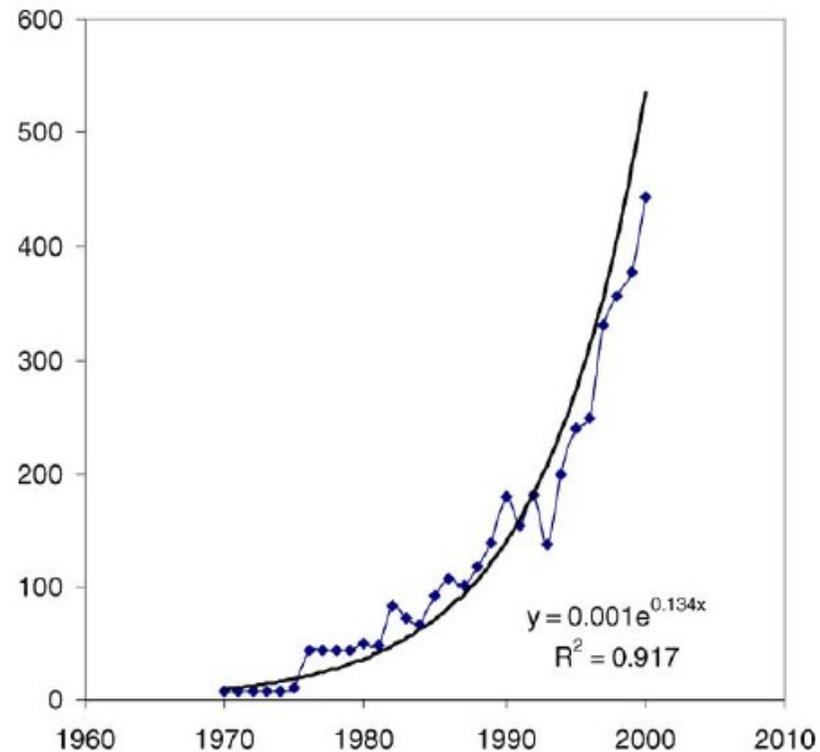


Figure 1. Exponential growth of publications indexed by Sociological Abstracts containing "social network" in the abstract or title.

## COME NASCE E A COSA SERVE LA SNA /5

### DIFFERENZA CON L'APPROCCIO TRADIZIONALE

	possiede le seguenti proprietà			
L'attore nella riga ...	Età	Sesso	Altezza	x ...
Marco				
Silvio				
Laura				
Dario				

	.. si relaziona con l'attore della colonna			
L'attore nella riga ...	Marco	Silvio	Laura	Dario
Marco	--	1	1	1
Silvio	0	--	0	1
Laura	1	0	--	1
Dario	1	0	0	--

#### Approccio convenzionale della ricerca sociale:

L'attenzione è sugli attributi dei singoli attori, il che porta di norma ad una matrice rettangolare.

Naturalmente, uno o più attributi possono essere proprietà di network, desunta dalla SNA.

E' una logica atomistica e compositiva.

#### Network Analysis:

L'attenzione è sulle relazioni fra attori, il che porta di norma ad una matrice quadrata.

Naturalmente, uno o più attributi degli attori possono essere presi in conto per esaminare le proprietà dei networks.

E' una logica strutturale, olistica e sistemica.

## TIPOLOGIE DI RETE ED OPERAZIONI SUI DATI/1

### TIPOLOGIE DI RETI

- **Una rete è un set di attori (nodi) uniti da un set di legami:**
  - gli attori possono essere persone, organizzazioni, concetti;
  - i legami possono essere diretti/indiretti; dicotomici o valorizzati;
  - si può vedere un rete come un insieme di reti (la più piccola delle quali è una diade), o una parte di una rete più ampia.
- Tipologie di reti (fra loro variamente combinabili):
  - **Reti costituite da legami orientati:** sono indicati da frecce che esprimono la direzione del legame. Rappresentano in genere comportamenti o giudizi degli attori (p.e. “A ama B”). Nel caso della *cognitive maps* esprimono causalità. Possono essere non simmetriche (e di solito infatti non lo sono, come nel caso “A ama B” e “B non ama A”).
  - **Reti costituite da legami vincolati:** sono indicati solo da linee senza frecce. Esprimono fatti in cui necessariamente gli attori implicati sono compresenti (p.e. “A comunica con B”). Sono simmetriche, ma possono essere altrimenti percepite.
  - **Reti a legame singolo:** esprimono un solo tipo di relazione fra attori (p.e. “A ama B”, oppure “A comunica con B”).
  - **Reti a legame multiplo:** esprimono contemporaneamente due o più tipologie di legami che uniscono gli attori (p.e. “A ama B” e “A comunica con B”).

## TIPOLOGIE DI RETE ED OPERAZIONI SUI DATI /2

### TIPOLOGIE DI RETI

- **Reti monomodali:** sono reti omogenee dal punto di vista degli attori che le costituiscono, cioè ognuno appartiene allo stesso tipo (p.e. “A” e “B” sono entrambi individui). Attenti a fare distinzione fra “tipo di attori” (p.e. individui, eventi, organizzazioni, ...) e valore degli attributi degli attori (p.e. individui con diverso sesso, livello di istruzione, ...).
- **Reti bimodali** (*two-mode networks*): sono reti che presentano contemporaneamente informazioni relative a due tipi di attori sociali (p.e. le singole persone e le organizzazioni a cui le stesse appartengono). Sono poco diffuse e un po' complesse.
- **Reti complete:** contengono tutti gli attori e tutte le loro relazioni (nei limiti di ciò che si intende analizzare).
- **Ego network:** è la porzione di rete di un particolare attore, vista dalla sua “soggettiva”.

## TIPOLOGIE DI RETE ED OPERAZIONI SUI DATI /3

### SCALE DI MISURA

- **Misura binaria:** è il livello più semplice e più diffuso. Convenzionalmente si indica con 1 la presenza di una relazione e con 0 la sua assenza. Essendo il “grado zero” dell'informazione, può sempre essere ottenuta a partire da un'altra scala di misura, definendo un valore di soglia (cut point) sotto il quale tutti i valori sono riportati a 0 e sopra ad 1. L'informazione che così si perde è spesso compensata dalla maggior facilità dell'analisi.
- **Misura su scala nominale a categoria multipla.** Per ogni relazione è indicato il tipo che essa assume, facendo riferimento ad un elenco a scelta multipla (esempio: amante, amico, collega, nemico, ...). L'analisi può essere svolta a livello di singolo tipo (p.e. networks che hanno “amante” come legame fra i nodi), con effetti sulle misure (p.e. riduzione della densità) di cui è importante avere coscienza. Come sempre può essere utile ridurre l'espressività alla misura binaria.
- **Misure ordinali.** La più semplice misura ordinale è riferita ad una scala a tre valori, del tipo “-1 0 +1”, dove il primo termine implica presenza di una relazione “negativa” (p.e. “avversione di un attore verso un altro”), il termine neutro l'indifferenza ed il termine positivo la situazione simmetrica a quella negativa. Altre misure ordinali fanno riferimento a scale più estese, di tipo Likert-like o p.e. basate sulla richiesta ad ogni attore di esprimere l'ordine (appunto) con cui gradirebbe avere relazioni con gli altri nodi del network. Gli ordinali possono sempre essere riportati ad una delle scale precedenti, perdendo informazione.
- **Misure scalari.** Le misure scalari sono state sviluppate più di recente, attraverso l'adattamento degli algoritmi originariamente nati per le ben più semplici misure binarie. Esse accompagnano spesso l'uso di tecniche di SNA al di fuori dell'ambito strettamente sociale, ad esempio in ambito economico o geografico. Nulla vieta di degradare l'informazione alle scale precedenti, stabilendo i cut point sulla base della natura del problema in esame.

## TIPOLOGIE DI RETE ED OPERAZIONI SUI DATI /4

### MODALITÀ DI RAPPRESENTAZIONE DEI DATI

- **Matrice di adiacenza.** E' una matrice quadrata  $n \times n$ , dove  $n$  è il numero degli attori che costituiscono il network in esame. E' l'usuale modalità di rappresentazione di una relazione fra attori. Per convenzione, le righe (i) contengono gli attori da cui la relazione “parte” e le colonne (j) gli attori a cui la relazione è rivolta. Una relazione è simmetrica quando  $X_{i,j} = X_{j,i}$ . Normalmente la diagonale principale della matrice è vuota e non è resa in conto nelle computazioni, ma – come si vedrà – in alcuni casi essa assume invece grande rilevanza.

	.. si relaziona con l'attore della colonna			
L'attore nella riga ...	Marco	Silvia	Dario	Laura
Marco	--	1	1	1
Silvia	0	--	0	1
Dario	1	0	--	1
Laura	1	0	0	--

- **Strati di matrici di adiacenza.** Nel caso in cui si hanno diversi tipi di relazioni, le stesse sono rappresentate ognuna attraverso una matrice di adiacenza – ognuna con il medesimo ordinamento delle righe e delle colonne – a loro volta “sovrapposte” in modo da creare una struttura “a strati”, detta “**multiplex**” (*multiple relations*). Un caso particolare di matrici multiple è costituito dalla matrici **CSS – Cognitive Social Structure**. Esse registrano le percezioni che i singoli attori hanno delle relazioni con gli altri nodi della rete (uno strato per attore).

## TIPOLOGIE DI RETE ED OPERAZIONI SUI DATI /5

### MODALITÀ DI RAPPRESENTAZIONE DEI DATI

- **Matrici multivaloriali.** Un'altra modalità che in alcuni casi è possibile utilizzare per rappresentare in una matrice di adiacenza più tipologie di legami è esprimere le relazioni in termini di variabili nominali politomiche, cioè che assumono diversi valori riferendosi a legami fra loro mutuamente escludenti o cumulativi. Ad esempio, indicando con “1.” una relazione di amicizia, con “2” una relazione propriamente matrimoniale, con “3” una relazione di amicizia e matrimoniale, etc. I valori sono puramente delle etichette di una variabile nominale, dunque non vanno confusi con una variabile ordinale o scalare. Vi sono alcuni algoritmi di analisi categoriale in grado di lavorare direttamente su questi dati, anche se conviene operare trasformazioni che riportino alla più usuale situazione delle matrici a strati.
- **Matrice degli attributi.** E' una matrice, spesso rettangolare, che riporta per ogni attore i suoi specifici attributi. Ai fini della *network analysis*, può svolgere diverse funzioni: *a)* esaminare le relazioni fra attributi e caratteristiche di rete; *b)* ospitare gli esiti di computazioni sulla rete; *c)* operare operazioni sulla rete quali la partizione e le permutazioni
- **Matrici di reti bimodali.** Vi sono due modalità tipiche di memorizzazione dei dati. La prima utilizza una matrice rettangolare costituita da  $n$  righe per  $m$  colonne, dove di solito  $n$  è il numero degli attori e  $m$  quello degli eventi a cui essi sono legati. Possono contenere utilmente dati di tipo ordinale (del tipo -1; 0; 1), rivolti ad indicare la presenza di una relazione positiva/negativa o la sua assenza. Come si vedrà oltre, di solito tali matrici sono rielaborate in matrici quadrate (p.e. “attori/attori” o “eventi/eventi”), in modo da essere esaminate secondo le usuali tecniche della SNA. Una seconda modalità utilizza **matrici bipartite**, ovvero formate ridondando righe e colonne (cioè aggiungendo alle righe le colonne ed alle colonne le righe), e codificando a 0 le intersezioni autoreferenziali. Si ottiene così una matrice quadrata, che consente l'applicazione delle usuali tecniche di analisi, ma prestando particolare attenzione all'interpretazione delle statistiche.

## TIPOLOGIE DI RETE ED OPERAZIONI SUI DATI /6

### MODALITÀ DI RAPPRESENTAZIONE DEI DATI

- Le **rappresentazioni grafiche** sono uno strumento particolarmente potente di analisi, basato sulle “proprietà intuitive” delle immagini, opportunamente guidate e migliorate dall'applicazione di misure statistiche in fase di scelta dei dati (nel caso di reti complesse) e dei layout (disposizioni dei nodi e dei legami nello spazio).
- La rappresentazione più semplice è il **sociogramma**, ovvero i nodi e le connessioni espresse dalla misura binaria (cioè presenti). In genere i sociogrammi sono resi con una disposizione casuale dei nodi, su cui è però possibile (e utile) intervenire per mettere in evidenza proprietà topologiche altrimenti invisibili. Una disposizione che facilita un primo approccio visivo ai dati è quella “**circolare**”, in cui tutti i nodi sono disposti simmetricamente lungo un'ellisse, in modo da massimizzare la loro distanza relativa, a fini di miglior lettura dei legami. Naturalmente le distanze relative non hanno qui significato analitico (spazio arbitrario).
- È possibile ed utile rappresentare le eventuali **differenze di attributi dei nodi** attraverso caratteristiche percettive quali il colore, la forma e la dimensione, potendo così essere trattate misure nominali, ordinali o scalari. Naturalmente fra gli attributi possono essere compresi esiti di misure di statistica delle reti (p.e. numero delle relazioni che partono da o che arrivano a un nodo, ...), in modo da rafforzarne l'interpretazione.
- Anche le **relazioni** possono essere oggetto di specifica visualizzazione, sia nel caso che esse siano espressione di misure nominali a scala multipla, ordinali o scalari, sia in modo da indicare proprietà comuni dei due nodi (p.e. utilizzare un colore per indicare la relazione fra due nodi che hanno in comune l'attributo “sesso = M”, un altro per l'attributo comune “sesso = F” ed un terzo nel caso eterosessuale).
- Sono inoltre usualmente applicabili in fase di visualizzazione operatori matematici (maggiore/minore) e booleani sul valore della relazione, in modo da effettuare delle **selezioni**. È inoltre possibile, nel caso in cui si disponga di **matrici multiplex**, rappresentare singolarmente o cumulativamente i diversi tipi di relazione, anche in modo differenziato.

## TIPOLOGIE DI RETE ED OPERAZIONI SUI DATI /7

### MODALITÀ DI RAPPRESENTAZIONE DEI DATI

- Una modalità sofisticata di rappresentazione è il **MDS – Multi Dimensional Scaling**, che dispone la rete nello spazio X-Y in modo che le loro distanze esprimano quanto i nodi sono fra loro “simili” o meno. In questo caso, le distanze assumono dunque un significato, che richiede comunque di essere accuratamente interpretato. (Per altri cenni al *multiscaling* vedi oltre).
- A volte, può essere utile semplificare volontariamente la struttura della rete, in modo da renderla più facilmente interpretabile in via grafica. Ad esempio, è possibile scegliere di eliminare i **nodi “isolati”** (cioè privi di legami) e i **nodi “pendenti”** (cioè uniti alla rete da un solo legame). Inoltre, possono essere visualizzate in modo specifico parti di rete che godono della proprietà di essere “sub-gruppi”, ovvero di presentare nodi dotati di caratteristiche simili in termini di struttura (componenti, blocchi, k-cores, fazioni, cluster, ..., vedi oltre).
- Infine, è sempre possibile rappresentare una rete dal punto di vista dei singoli attori, cioè in termini delle diverse **ego-net** che la costituiscono, ovviamente più semplici – e dunque più leggibili visivamente – del network complessivo, soprattutto se questo si presenta molto ricco di attori e legami.
- **Non esiste a priori un modo “giusto” di rappresentare un network:** bisogna provare diverse configurazioni, avendo chiaro cosa si vuole mettere in evidenza e perché (cioè il significato interpretativo della rappresentazione). Vedi per un approfondimento la lettura “McGrath, C., Krackhardt D., Blythe J., Visualizing Complexity in Networks: Seeing Both the Forest and the Trees, *Connections* 25, 2003”.

## TIPOLOGIE DI RETE ED OPERAZIONI SUI DATI /8

### OPERAZIONI SULLE MATRICI

- Non sempre le strutture dei dati si presentano coerenti con il tipo di elaborazioni da svolgere. Il che implica di operare delle trasformazioni. Di seguito si presentano quelle più usuali, facendo riferimento alle funzioni del programma Ucinet.
- **Permutazioni.** Hanno il compito di mettere in maggior evidenza le relazioni fra i dati, modificando l'ordine delle colonne della matrice, in modo da creare utili prossimità fra nodi, dotati p.e. di attributi uguali. Normalmente ( e sicuramente nel caso in cui la matrice è simmetrica), la permutazione delle colonne è attuata in modo da portare ad una analoga permutazione delle righe, in modo da mantenere la coerenza interna.
- **Partizione.** E' un'operazione di suddivisione di una matrice in blocchi, cioè in reti di dimensione minore (o super nodi), caratterizzate p.e. da uniformità di attributi fra gli attori o da altro criterio di loro identificazione. L'esito di una partizione è una matrice di adiacenza in cui ai nodi originari (attori) sono sostituiti i super nodi. Non si confonda l'operazione di partizione (la cui *ratio* è definibile a piacere dall'analista) con la clusterizzazione, definita attraverso applicazione di algoritmi ad hoc. La partizione è una tecnica utilizzata per sviluppare ipotesi sui ruoli sociali determinati da attributi, oggetto di verifica attraverso esame comparativo di misure effettuate su ogni blocco risultante, fra cui in particolare il calcolo di:
  - **densità di blocco**, intesa come il rapporto fra il numero dei legami effettivi fra i nodi costituenti il blocco ed il numero dei legami teorici dello stesso, senza tenere conto della diagonale principale (vedi oltre);
  - **immagine del blocco**, intesa come dicotomizzazione delle densità di blocco, attraverso la scelta di un opportuno *cut point*.

## TIPOLOGIE DI RETE ED OPERAZIONI SUI DATI /9

### OPERAZIONI SULLE MATRICI

- **Collassamento.** E' un'operazione che porta a “fondere” due o più righe/colonne in un'unica riga/colonna, attraverso l'applicazione di operatori quali somma, media, valore massimo o minimo. Ha lo scopo di semplificare la matrice, riducendo il numero dei nodi. Come nel caso della partizione, la *ratio* è definibile a piacere dall'analista.
- **Trasposizione.** In analisi delle reti, la trasposta di una matrice di adiacenza esprime la sorgente dei legami posseduti da un attore. Il grado di similarità fra una matrice e la sua trasposta esprime in modo sintetico il grado di simmetria delle relazioni fra attori.
- **Dicotomizzazione.** E' l'operazione che consente di ricondurre qualunque scala ad una misura binaria, attraverso la fissazione di un *cut point*. E' propedeutica all'applicazione di algoritmi utilizzabili solo su matrici binarie. Consente inoltre di generare, in funzione del *cut point*, diverse matrici da quella originaria, p.e. ognuna rappresentativa di un “livello” di analisi (p.e. “strettezza delle relazioni”, “distanza fra gli attori”, etc.).
- **Simmetrizzazione.** E' l'operazione che rende “forzatamente” simmetrica una matrice di adiacenza, applicando un criterio di attribuzione di valore ad uno o ad entrambi i valori interessati. Funzioni tipiche sono: il massimo ( $ij = \max ji$  o viceversa), il minimo, la media, la somma, la differenza, il prodotto, la divisione. Un'altra modalità di interesse è la trasformazione speculare lungo la diagonale principale, copiando la “metà bassa” della matrice in quella “alta” o viceversa: ciò consente di focalizzare l'attenzione rispettivamente sui nodi che inviano o sui nodi che ricevano legami da/verso altri nodi. Naturalmente, le tecniche di simmetrizzazione – al di là dei loro aspetti matematici – vanno scelte in funzione del significato sociologico (o comunque analitico) che presentano.

## TIPOLOGIE DI RETE ED OPERAZIONI SUI DATI /10

### OPERAZIONI SULLE MATRICI

- **Normalizzazione.** E' un'operazione che consente di esprimere i valori della matrice in forma percentuale, con riferimento ai totali di riga, di colonna, alla somma dei valori della matrice, alle medie di riga/colonna al valore massimo di un elemento di riga/colonna o alle distanze euclidee. Normalmente va applicata su misure scalari, e senz'altro non su ordinali. Anche in questo caso, si tratta di avere ben chiaro il significato sociologico (o comunque analitico) delle operazioni sui dati, visto che a seconda dell'algoritmo di normalizzazione di ottengono esiti molto diversi.
- **Trasformazione di attributi in relazioni.** Normalmente, come si è visto, una rete è costituita da relazioni fra attori, ognuno dei quali dotato di attributi. Si può anche immaginare di “rovesciare” la rete, ridefinendo le relazioni a partire dal rapporto fra gli attributi degli attori. Se, ad esempio, due attori condividono lo stesso attributo (p.e. hanno entrambi sex=“M”), la relazione fra di essi assume il valore 1 (altrimenti 0). Esistono differenti funzioni di confronto fra attributi (somma, prodotto, differenza, differenza assoluta, ...), che vanno scelte in ragione del senso dell'analisi.
- **Affiliazione.** Consente di trasformare una rete bimodale (p.e. “attori/eventi”) in una rete mono modale, a scelta p.e. fra “attori/attori” o “eventi/eventi”, sulla base delle proprietà delle relazioni che li accomuna all'altro tipo di nodo (rispettivamente nell'esempio “eventi” e “attori”). Per i legami binari, il metodo usuale è il *cross-product*, ovvero la somma dei prodotti binari degli attori rispetto agli eventi (che equivale a dire i casi in cui vi è co-occorrenza) . Per i legami valorizzati, il metodo usuale è la somma del valore minimo del legame fra attore ed eventi.

## TIPOLOGIE DI RETE ED OPERAZIONI SUI DATI /11

### OPERAZIONI SULLE MATRICI

- **Bipartizione.** Sempre in tema di matrici bimodali, un'altra modalità di loro riduzione ad una matrice monomodale è data dalla creazione di una struttura ridondante, aggiungendo alle righe (p.e. gli attori) le colonne (ovvero p.e. gli eventi) e simmetricamente aggiungendo alle colonne le righe. Su questa nuova matrice sono rappresentate le relazioni originarie.
- **Trasformazione di relazioni in attributi.** Invece di dire che una rete è un insieme di attori connessi da relazioni, possiamo sempre dire (forse in un modo un po' barocco) che una rete è *“un insieme di relazioni connesse da attori”*. Il che ci porta a trasformare la matrice di adiacenza originaria in una nuova matrice che contiene nelle righe gli attori e nelle colonne le diverse relazioni effettive da cui la rete è costituita. Tale nuova **matrice di incidenza** (detta anche **ipergrafo**) può prevedere il ricorso ad una scala di misura ordinale, di solito del tipo “-1, 0, 1”, dove il segno positivo indica che il legame parte dal nodo ed il segno negativo che è ricevuto dal nodo.
- **Trasformazione linegraph.** Un'ulteriore rilettura non ovvia porta a rappresentare una rete come *“relazioni fra relazioni”*, ovvero una matrice in cui sia le righe che le colonne sono costituite da relazioni (cioè da coppie di nodi per cui vi sia un legame), i legami stessi sussistendo solo quando le stesse sono connesse attraverso un nodo nel grafo originario.
- **Trasformazione multiplex.** Consente di trasformare una matrice di adiacenza a strati in un'unica matrice multivaleoriale, assegnando automaticamente ad ogni tipo di relazione riscontrata un valore. Si ricordi che detto valore è espressione di una variabile nominale politomica, non ordinale o scalare !!!

## TIPOLOGIE DI RETE ED OPERAZIONI SUI DATI /12

### OPERAZIONI SULLE MATRICI

- **Trasformazione multigraph.** E' l'inverso della precedente, ovvero consente di trasformare una matrice multivaloriale in un insieme di strati di matrici di adiacenza binarie, una per ogni valore della relazione, lo stesso visto come variabile nominale politomica. Ovviamente, se si applicasse multigraph all'esito di una precedente trasformazione multiplex non si tornerebbe alla situazione originale.
- **Funzioni di semplificazione.** Se si hanno reti di dimensioni rilevanti, può essere utile cercare di semplificarle, in modo da metterne in risalto gli aspetti chiave, senza ovviamente alterarne la struttura. Nel tempo sono state sviluppati in modo *bottom-up* diversi approcci, a cui ricorrere anche “tentativamente”. In particolare:
  - **Estrazione della componente principale.** Estrae i nodi e le loro relazioni che nel complesso costituiscono il sub-grafo di maggior dimensione, ovvero hanno il maggior numero di vertici.
  - **Rimozione dei pendenti e degli isolati.** Elimina i nodi che non hanno legami (isolati) o che sono connessi al grafo attraverso un solo legame (pendenti).
  - **Estrazione di Egonet.** Estrae una nuova matrice costituita dai soli legami diretti di uno o più nodi a scelta (ego).

## MISURE DI RETE /1

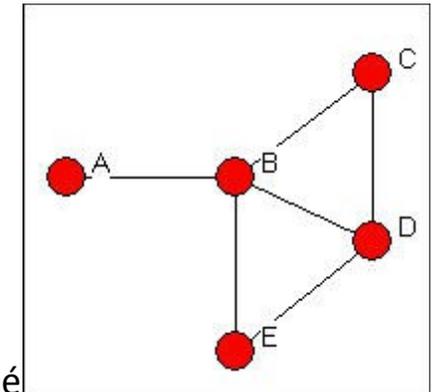
### MISURE GENERALI E DI CONNESSIONE

- **Ampiezza.** L'ampiezza di una rete è espressa in termini di numero dei nodi presenti. In generale tale valore non ha un particolare significato analitico, fatto salvo che maggiore è l'ampiezza, maggiore è la probabilità di ogni singolo attore di avere una visione “limitata” della rete complessiva.
- **Grado (*degree*).** Il grado di un nodo è il numero complessivo dei legami che esso possiede, senza tener conto della loro direzione. Così espressa, questa misura è tipica di una rete a legami vincolati, ovvero non orientati. In caso contrario, è molto più utile misurare l'**in-degree** e l'**out-degree**, intesi come il numero dei legami che arrivano o partono da un dato nodo, espressi – per una matrice di adiacenza – rispettivamente dal totale di colonna o di riga. E' utile, a fini di confronto fra reti, normalizzare il valore assoluto rapportandolo al numero complessivo di legami possibili (cioè a  $k-1$ , dove  $k$  è il numero degli attori del network). Le misure di in e out-degree hanno grande importanza in termini sociologici, esprimendo il ruolo svolto dagli attori interessati. La loro interpretazione non può prescindere dal significato del tipo di legame in esame.
- **Densità.** Nel caso di una rete a legame binario la densità è il rapporto fra il numero dei legami effettivi ed il numero delle diadi (coppie), ovvero di tutti i possibili legami diadici teoricamente esistenti. Se i legami sono espressi con variabili scalari (p.e. forza o probabilità della relazione) la densità è espressa come rapporto fra la somma dei valori di tutti i legami e il numero dei legami possibili (ovvero è la “forza media” dei legami teorici).
- **Raggiungibilità.** Indica se un attore è raggiungibile da un altro, indipendentemente dalla lunghezza del percorso e tenendo in conto l'orientamento dei legami. Naturalmente in caso di matrici non simmetriche si può dare il caso che “A” sia raggiungibile da “B”, ma non il contrario. Sociologicamente, la raggiungibilità (che è una proprietà di connessione) va interpretata in ragione del significato del tipo di relazione che si osserva.

## MISURE DI RETE /2

### MISURE GENERALI E DI CONNESSIONE

- **Connettività.** E' una misura complementare della raggiungibilità, esprimendo il numero di nodi che dovrebbero essere cancellati dalla rete per disconnettere due attori. In generale, maggiore è il numero, maggiore è la possibilità dell'attore di ricevere/trasmettere informazioni; operare scambi, etc., soprattutto se i nodi sono dati dalla sommatoria di diversi percorsi. Minore è il numero, maggiore è probabilmente la vulnerabilità dell'attore interessato.
- **Distanza.** E' una misura della connessione fra due attori, e può essere definita facendo riferimento a tre diverse logiche di misura, a mano a mano più restrittive.
  - Nel caso delle reti vincolate (cioè in cui non vi è direzione dei legami):
    - **cammino (walk).** E' la più generale, ed esprime il numero dei passi che occorrono per passare da un nodo ad un altro, senza restrizioni. Ad esempio, da "A" a "C" vi sono: un cammino di lunghezza 2 (ABC); 1 cammino di lunghezza 3 (ABDC); 3 cammini di lunghezza 4 (ABEDC; ABDBC; ABEBC), etc.;
    - **tragitto (trail).** E' una restrizione di cammino, in modo da non utilizzare le relazioni più di una volta. Ad esempio, il numero dei cammini di lunghezza 4 si riduce ad 1 (ABEDC) e vi è un cammino di lunghezza 5 (ABDEBC);
    - **percorso (path).** E' una ulteriore restrizione di cammino, in modo da non utilizzare né le relazioni né i nodi più di una volta. Nell'esempio, restano i cammini di lunghezza 2 (ABC); 3 (ABDC) e 4 (ABEDC), ma scompare quello di lunghezza 5. Il percorso è l'usuale riferimento per misurare le distanze in una rete.
- Nel caso delle reti costituite da legami orientati, vale quanto fino a qui visto, tenendo conto della direzione del legame stesso. Sono in uso le denominazioni di "semi-walk", "semi-trail" e "semi-path".



## MISURE DI RETE /3

### MISURE GENERALI E DI CONNESSIONE

- **Distanza geodetica.** E' il percorso di minor lunghezza (maggiormente efficiente) che unisce due nodi. Nel caso di reti a legame binario, è pari semplicemente al numero dei legami che costituiscono il percorso. Nel caso di reti con legame valorizzato, si parla di “distanza generalizzata”, il che richiede preventivamente di interpretare il significato del legame. Si distinguono (qui, come in generale in SNA), tre situazioni:
  - il legame rappresenta un “**costo**” (e in generale una quantità di risorse economiche impegnate in una relazione, come p.e. nel caso del “costo di transazione”): la distanza geodetica è relativa al percorso meno “costoso”, cioè che minimizza la somma dei costi dei diversi percorsi possibili;
  - il legame è rappresentato in termini di **forza** della relazione, la stessa essendo per ogni percorso pari al valore del suo segmento più debole (il “principio della catena”): la distanza geodetica è il percorso che contiene i segmenti di maggior “forza” (ovvero che non contiene i segmenti di maggior debolezza);
  - il legame rappresenta una **probabilità**: la distanza geodetica è data dal percorso che massimizza il prodotto delle probabilità dei singoli legami, fra i diversi possibili.

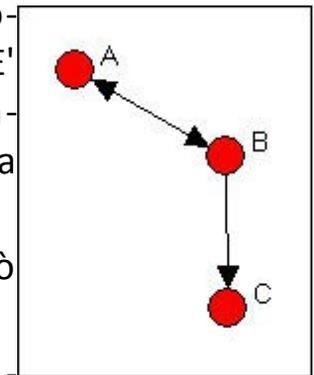
E' inoltre possibile esprimere la distanza come “vicinanza” o “**prossimità**”, il che però richiede di definire un termine di confronto, attraverso funzioni di tipo:

- **moltiplicativo**: il valore della prossimità è dato dal rapporto fra la maggior distanza teorica possibile e la distanza geodetica, ovvero  $(n. \text{ nodi} - 1) / \text{distanza nodo}$ ;
- **additivo**: il valore della prossimità è dato dalla differenza fra il numero dei nodi e la distanza geodetica;
- **lineare**: il valore è riscalato rispetto ad una scala, normalizzata fra 0 e 1, ovvero:  $1 - [(distanza \text{ nodo} - 1) / (n^\circ \text{ nodi} - 1)]$ ;
- **decadimento di frequenza**. La prossimità di due nodi è diminuita dalle distanze pari o minori che il nodo obiettivo (quello a cui Ego si rivolge) ha con altri nodi. La prossimità è  $1 -$  la proporzione degli attori che hanno con *alter* lo stesso (od un miglior) rapporto di Ego.

## MISURE DI RETE /4

### MISURE GENERALI E DI CONNESSIONE

- **Eccentricità e diametro.** L'eccentricità è la lunghezza del maggior percorso geodetico che lega Ego ad un altro nodo della rete. Il diametro invece è la lunghezza del maggior percorso geodetico che vi è in assoluto nella rete; tale valore assume spesso la funzione di defintore della profondità con cui la stessa è indagata, ovvero del numero dei passi di percorso che sono presi in considerazione nelle analisi.
- **Numero di percorsi geodetici.** Esprime di quanti percorsi alternativi di natura geodetica dispone ogni attore verso ogni altro attore, rappresentando una misura della ridondanza della rete. Maggiore il numero, minore probabilmente il potere di “mediazione” attribuito agli attori, visto che “ognuno appare variamente sostituibile”.
- **Flusso massimo.** Anche questa è una misura rivolta a comprendere la possibilità di un attore di comunicare con gli altri in una rete, interpretabile in termini di vulnerabilità o di ridondanza. E' espressa dal numero degli attori attraverso cui passa una connessione fra Ego ed Alter, indipendentemente dalla lunghezza del percorso. Maggiore il valore, maggiore in senso probabilistico la “connessione” di Ego, al di là della sua efficienza.
- **Reciprocità.** La reciprocità è intesa come simmetria della relazione all'interno di una diade, e può essere misurata in due modi differenti:
  - Il primo considera il rapporto fra le diadi con legame reciproco ed il totale delle diadi effettivamente presenti. Nel nostro caso la reciprocità vale 0,500.
  - il secondo il rapporto fra il numero di archi reciproci (che vanno contati come doppi) ed il numero totale degli archi effettivamente presenti. Nel nostro caso la reciprocità vale 0,667.



Normalmente la reciprocità delle diadi è la misura più utilizzata, in particolare in reti ampie. In sociologia, la reciprocità è vista da molti come una condizione fondante la stabilità di una diade.

## MISURE DI RETE /5

### MISURE GENERALI E DI CONNESSIONE

- **Transitività**. La transitività è una proprietà delle triadi, cioè di relazioni fra tre attori. Per molti sociologi, le triadi assumono un'importanza rilevante, in quanto danno (darebbero) conto di diversi “funzionamenti” strutturali. In senso tassonomico, si possono immaginare 4 configurazioni tipiche di tre attori:
  - **isolamento**, ovvero assenza di ogni tipo di legame fra ognuno di essi;
  - relazione di **diade**, ovvero solo due attori su tre hanno una relazione;
  - “**buco strutturale**”, ovvero due attori sono connessi ad un terzo, ma non lo sono direttamente fra di loro;
  - “**cluster**”, in cui i tre attori sono connessi gli uni agli altri a formare un “grappolo”, che può essere dotato di transitività, ovvero presentare le relazioni AB, BC, e AC.

Per un grafo binario ed orientato, la tipica misura della transitività è il conto delle adiacenze, cioè quante volte le due relazioni AB e BC sono accompagnate dalla relazione AC. Se il grafo ha legami valorizzati, si distingue fra:

- **transitività forte**: quando il legame AC è maggiore del valore di soglia massima, prefissato dal ricercatore;
- **transitività debole**: quando il legame AC è compreso fra la soglia minima e la soglia massima, entrambe sempre definite dal ricercatore, in funzione delle proprie esigenze di analisi.

Altre misure di transitività tipiche di grafi con legami valorizzati sono:

- **misura euclidea**: quando il legame AC è minore di  $AB + BC$ ;
- **misura stocastica**: quando il legame AC è minore di  $AB * BC$ .

In generale le misure di transitività sono normalizzate rapportandole a due possibili denominatori:

- il numero complessivo delle triadi di qualunque tipo presenti nel grafo;
- il numero delle triadi complessivamente costituite da AB e da BC, cioè il numero massimo di triadi potenzialmente transitive.

## MISURE DI RETE /6

### MISURE GENERALI E DI CONNESSIONE

- **Coefficiente di *clustering***. La possibilità di un attore di agire in una rete è, fra l'altro, molto legata alla densità delle relazioni dei nodi con cui lui stesso è legato. Se Ego è legato a pochi nodi, ma ognuno di essi è molto ricco di connessioni con il resto del grafo, può essere in una situazione migliore che se si trovasse legato a molti nodi “poveri” dal punto di vista delle loro ulteriori relazioni con il resto della rete. Questa differenza è misurata in termini di “coefficiente di *clustering*”, inteso come la densità della rete dei nodi con cui Ego ha relazioni. Il coefficiente può essere calcolato:
  - per i singoli nodi, come densità delle reti proprie dei nodi con cui è connesso (i nodi “vicini”);
  - per la rete complessiva, come media (semplice o pesata sulla base del *degree* di ogni attore) del coefficiente dei singoli nodi. E' importante riflettere sulla differenza fra tale misura e il valore della densità della rete (vedi prima): il coefficiente di *clustering* complessivo è di solito maggiore della densità, esprimendo in tal modo la “strana proprietà” delle reti di consentire mediamente una elevata possibilità di relazione fra attori anche in presenza di basse densità (→ vedi *small words*).
- **External-Internal index**. Sempre in tema di analisi delle proprietà di connessione di un network, una misura possibile solo nel caso in cui si disponga di partizioni è l'E-I index, dato per ogni gruppo dalla differenza fra il numero dei legami che esso ha con attori esterni ed il numero dei legami interni allo stesso, diviso il numero totale dei legami senza tenere conto della direzione del legame. L'indice E-I ha un campo di esistenza teoricamente compreso fra -1 (tutti i legami sono solo interni) e + 1 (il contrario). Oltre che a livello di gruppo, l'indice è calcolabile con riferimento all'intero network e per ogni singolo attore. Va testata la significatività della correlazione fra valore dell'indice e gruppo, confrontandolo con i valori calcolati su matrici definite permutando in modo casuale i legami e mantenendo invariata la composizione dei blocchi ed il valore complessivo della densità.

## MISURE DI RETE /7

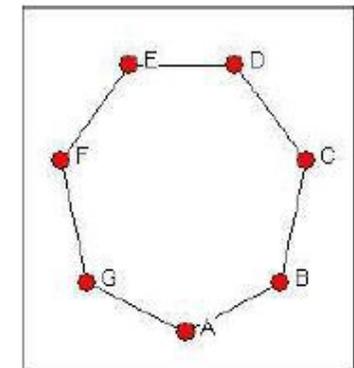
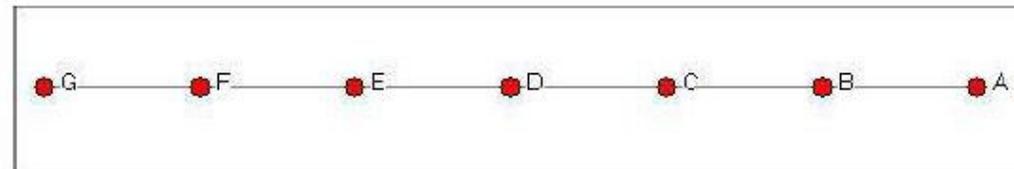
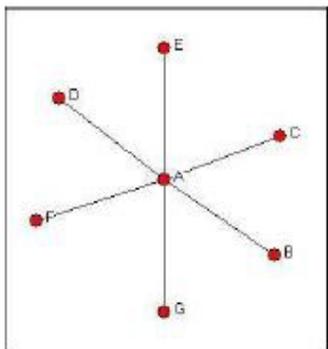
### MISURE GENERALI E DI CONNESSIONE

- **Indici di gerarchia di Krackhardt.** Sono indici costruiti assumendo come riferimento un modello organizzativo di totale natura gerarchica (visto come una “piramide” caratterizzata da totale asimmetria delle relazioni e unicità del comando) e comparando ad esso il network in esame, al fine di rivelare i suoi “scostamenti” da tale forma idealtipica. Il loro utilizzo non è esente da critiche, e dunque richiede (come sempre) discernimento. Gli indici sono:
  - **Connettività:** in una gerarchia pura, la rete deve avere un solo componente (cioè tutti i nodi sono parte della stessa struttura, visto che ve ne è sicuramente uno – il “boss” – a cui tutti gli altri fanno per forza riferimento). Si misura la distanza da questo riferimento guardando al rapporto fra il numero di coppie raggiungibili ed il numero di coppie totali, ovvero alla proporzione degli attori non raggiungibili dagli altri. Nell'organizzazione gerarchica (ma non solo in essa), tale indice è eguale a 1;
  - **Gerarchia:** come si è detto, in una gerarchia pura non possono esistere legami reciproci, visto che essi implicherebbero un eguale status sociale fra gli attori implicati. Si misura il rapporto fra il numero delle coppie che non hanno rapporti reciproci ed il numero totale delle coppie. Nell'idealtipo tale indice è eguale a 1;
  - **Efficienza:** come detto, in una gerarchia pura ognuno (tolto il “boss supremo”) ha uno ed un solo capo, ovvero – in termini di reti – ha un in-degree pari a 1. Nell'idealtipo tale indice è uguale a 1.
  - **“LUBness” (Least Upper Boundness):** in una gerarchia pura il comando è unificato, cioè ogni coppia di attori (eccettuate le coppie formate fra il boss supremo e gli altri membri) deve avere un legame con lo stesso attore sovraordinato. Si misura come rapporto fra il numero di coppie che non hanno un nodo comune ed il numero delle coppie che potrebbero averlo. Nell'idealtipo tale indice è uguale a 1.

## MISURE DI RETE /8

### MISURE DI CENTRALITÀ

- Uno specifico insieme di misure è relativo alla “centralità” di ogni attore, la stessa potendo essere assunta (con cautele) come indicatore del potere. Dal punto di vista empirico, possiamo immaginare tre caratteristiche posizionali di un attore a questo fine rilevanti:
  - il **grado**, come misura di connessione: maggiore è il grado, maggiore è probabilmente il potere dell'attore, in quanto dispone di maggior libertà nella scelta d'uso dei propri legami o se si preferisce, è meno dipendente dagli altri;
  - la **prossimità**, come misura di distanza (vicinanza) dagli altri attori: minore è la distanza (espressa p.e. in termini di lunghezza dei percorsi geodetici), maggiore può essere il potere derivante dall'essere un “punto di riferimento” per gli altri attori, poterli raggiungere facilmente, etc.
  - la **betweenness**, come misura del ruolo di connettore di altri attori, assumendo una funzione di broker: maggiore la betweenness, maggiore è probabilmente il potere posseduto.
- Tre tipologie idealtipiche di network aiutano a comprendere il funzionamento delle caratteristiche sopra esposte: la rete a **stella**, la rete **lineare** e la rete **circolare**.



## MISURE DI RETE /9

### MISURE DI CENTRALITÀ

- Nella **rete a stella**, l'attore A gode del massimo valore di grado (6), della minor distanza geodetica da tutti gli altri attori (1) e della maggior betweenness, visto che si presenta comune ad ogni coppia di altri attori. Tutti gli altri nodi si presentano eguali dal punto di vista degli attributi di potere ora esaminati. La stella è la forma di rete di minor eguaglianza.
- Nella **rete lineare**, la situazione è un po' più complessa. Gli attori A e G sono svantaggiati dal punto di vista del grado, della vicinanza e della betweenness, mentre gli altri si trovano in una situazione abbastanza equivalente (tolto il caso di D, che è migliore sotto il profilo della distanza geodetica).
- Nella **rete circolare**, ogni attore ha gli stessi attributi di grado, vicinanza e betweenness. Il circolo è la forma di rete di maggior eguaglianza.

### MISURE DI PROSSIMITÀ BASATE SUL GRADO

- In generale la **centrality degree** può essere espressa per ogni attore come valore dell'indegree e dell'outdegree posseduto (nel caso ovviamente di reti a legami orientati), usualmente normalizzato come rapporto percentuale sul numero degli attori meno uno. Rilevanti sono inoltre le misure a livello di network complessivo (valor medio e relative caratteristiche della distribuzione dei valori che lo formano), che consentono di valutare quanto il network si presenta connesso e (dis)omogeneo.

## MISURE DI RETE /10

### MISURE DI CENTRALITÀ

Sono disponibili misure più sofisticate:

- **misura di centralità di Freeman**: esprime come valore percentuale quanto il network sia simile alla rete idealtipica a stella (cioè al massimo della distribuzione diseguale del potere): maggiore il valore, maggiore la concentrazione di potere;
- **misura di centralità di Bonacich**: è una misura basata sull'idea che non vi è necessariamente una relazione lineare fra centralità e potere. Da un lato, la centralità di Ego è maggiore se i nodi a cui è connesso hanno a loro volta rilevanti connessioni con il resto della rete. Dall'altro, Ego ha più potere se i nodi a cui è connesso hanno poche relazioni con il resto del network, ovvero sono più dipendenti da lui stesso. L'algoritmo iterativo di Bonacich permette di valutare entrambe le situazioni, lasciando al ricercatore la possibilità di definire segno ed intensità del capitale relazionale degli attori cui Ego è connesso.

### MISURE DI CENTRALITÀ BASATE SULLA PROSSIMITÀ

- Differentemente dalle misure basate sulla centralità, che al più prendono in conto solo le relazioni dei nodi con cui Ego è direttamente connesso, quelle riferite alla prossimità si riferiscono all'intero insieme delle relazioni che Ego ha con tutti gli attori (raggiungibili) del network.

## MISURE DI RETE /11

### MISURE DI CENTRALITÀ

- Fra quelle più usuali ricordiamo:
  - la “lontananza” (**far-ness centrality**) ed il suo reciproco (**closeness centrality**). La prima è data dalla somma delle lunghezze dei percorsi geodetici che legano Ego a tutti gli attori cui è connesso. La seconda è il reciproco della prima (1/far-ness), normalizzato dividendolo per la far-ness minima, ed esprimendo il risultato in percentuale;
  - la **reach centrality**, misurata come numero dei nodi che Ego può raggiungere in 1, 2 3 o n passi, normalizzato dividendolo per il massimo valore della raggiungibilità;
  - l'**eigenvector (autovettore) della distanza geodetica**, che esprime la centralità di un attore rispetto alla prima dimensione (componente principale) derivata dall'analisi fattoriale delle distanze fra gli attori costituenti la rete . E' una misura complementare alla closeness centrality, differenziandosene per la maggior rappresentazione dell'intera struttura del network, La sua interpretazione richiede preliminarmente di valutare la “bontà di rappresentazione” della componente principale dell'analisi fattoriale, che deve spiegare da sola almeno il 75% della struttura della rete (ovvero la struttura complessiva) Attualmente il calcolo dell'eigenvector è limitato a matrici simmetriche, prendendo dunque in conto solo la “presenza” di un legame, invece che la sua direzione.
  - **misure di coesione di Hubbel, Katz e Taylor**. E' una misura della forza dell'insieme di tutte le relazioni che legano due attori all'interno di un network, definita pesando il contributo di ogni singola relazione in termini inversamente proporzionali alla lunghezza del percorso. Le prime due sono riferite ad una matrice simmetrica e sono espresse con riferimento ad una scala compresa fra il minimo ed il massimo della distanza fra gli attori dell'intero network. La terza, analoga alle precedenti, è relativa a reti costituite da legami orientati, ed è data dalla differenza fra l'out degree e l'in degree, normalizzata sul numero dei passi del percorso che lega i due attori.

## MISURE DI RETE /12

### MISURE DI CENTRALITÀ

#### MISURE DI CENTRALITÀ BASATE SULLA BETWEENNESS

- Le misure basate sulla betweenness esprimono la posizione di intermediazione fra due nodi svolta da un attore in quanto membro di un percorso geodetico, rapportata al numero dei percorsi geodetici che connettono i due nodi. Se vi è un solo percorso geodetico, il valore di betweenness sarà elevato, in quanto il mediatore non sarà “by-passabile” dai due nodi che contribuisce a connettere; se gli stessi sono uniti da più percorsi geodetici, il potere del nostro attore sarà decrescente in funzione del numero degli stessi. Per un network con legame binario si distinguono due misure:
  - **betweenness centrality di Freeman, a livello di nodi**. Per ogni attore, è l'inverso del rapporto fra il numero dei percorsi geodetici che legano ogni coppia di nodi e il numero delle volte in cui ricade in essi ogni singolo attore, normalizzato rispetto alla betweenness massima teorica;
  - **betweenness centrality di Freeman, a livello di legami**. E' analoga alla precedente, ma mette in evidenza la centralità dei singoli legami che uniscono gli attori;
  - **flow centrality**. E' una generalizzazione della misura di Freeman a livello di nodi, relativa a tutti i percorsi che uniscono due attori, invece che solo a quelli geodetici, pesati in ragione inversa della loro lunghezza. E' di solito normalizzata per singolo attore rapportandola al valore complessivo della centralità

## MISURE DI RETE /13

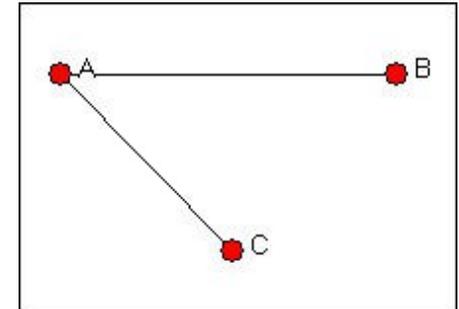
### MISURE RELATIVE AD EGO

- **Misure di connessione.** Sono in larga parte equivalenti alle principali misure di connessione già incontrate per il network complessivo, e sono in genere riferite alla sola porzione di rete costituita dagli attori cui Ego è legato (cioè dagli attori denominati “Alter”), non tenendo in conto i legami fra Ego ed Alter. Le misure possono essere riferite al network costituito da: *i)* i nodi che inviano relazioni a Ego; *ii)* i nodi a cui Ego invia relazioni; *iii)* entrambi i nodi.
  - **Densità:** numero dei legami fra gli attori Alter, rapportato al numero dei legami teoricamente possibili ed espresso in valore percentuale;
  - **Distanza geodetica media:** distanza calcolata con riferimenti ai legami fra gli attori Alter, nel caso in cui tutti siano fra loro raggiungibili (anche attraverso Ego);
  - **Diametro:** il percorso geodetico di maggior lunghezza nel network di Ego;
  - **Numero dei componenti deboli:** è il numero dei subgrafi che sono fra loro connessi solo attraverso Ego, indipendentemente dal verso della relazione (ovvero dei subgrafi che resterebbero fra loro disconnessi se Ego fosse rimosso);
  - **Efficienza della raggiungibilità:** numero degli Alter raggiungibili da Ego in 2 passi, normalizzato come percentuale del numero totale degli Alter più la somma della dimensione dei loro propri networks.
  - **Brokerage:** numero delle coppie di Alter connesse attraverso Ego, normalizzato rispetto al numero delle coppie di Alter.
  - **Betweenness:** inverso del rapporto fra il numero dei percorsi geodetici che legano ogni coppia di nodi Alter e il numero delle volte in cui Ego ricade in essi, normalizzato rispetto al numero massimo dei percorsi geodetici.

## MISURE DI RETE /14

### MISURE RELATIVE AD EGO

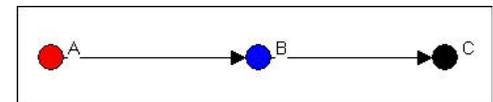
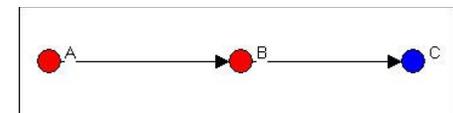
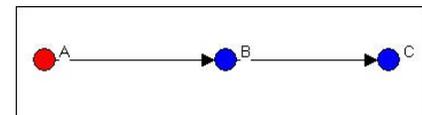
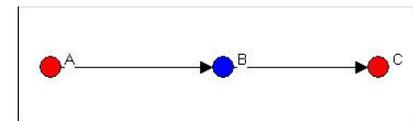
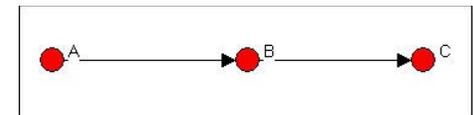
- **Buchi strutturali.** Sono una espressione del potere di Ego (A), visto come “mediatore necessario” delle relazioni fra altri due nodi (B, C). Secondo Burt sono possibili diverse misure abbastanza complesse, tutte rivolte a cogliere il rapporto fra Ego e la struttura del suo network personale, nelle non intuitive relazioni fra “potere” e “vincolo”. Fra quelle più usuali segnaliamo:
  - **ridondanza diadica:** esprime attraverso una matrice di adiacenza, per ogni Alter, quanti attori oltre ad Ego svolgono un ruolo di “mediazione” verso gli altri Alter. Maggiore la ridondanza, minore la dimensione del buco strutturale e ovviamente minore il potere di Ego verso l'attore per cui la stessa è misurata;
  - **vincolo diadico:** esprime attraverso una matrice di adiacenza quanto gli Alter su cui Ego ha “investito” (attivando relazioni) hanno a loro volta investito in attori diversi da Ego. Maggiore è il valore del vincolo diadico, minore è il “potere” di Ego;
  - **dimensione effettiva della rete:** è il numero degli Alter di Ego diminuito del grado medio degli Alter all'interno del network di Ego, senza prendere in conto i legami con Ego. Maggiore la dimensione, maggiore il “potere” di Ego;
  - **efficienza:** il valore precedente, diviso per il numero degli Alter che costituiscono la rete di Ego;
  - **vincolo:** è una misura sostanzialmente analoga alla ridondanza diadica, ma di carattere sintetico, ovvero non dettagliato per tutte le coppie di reazioni fa Ego ed i singoli Alter. Maggiore è il valore del vincolo, maggiore è il “condizionamento” che gli Alte possono esercitare su Ego, in quanto in grado di disintermediarsi da esso;
  - **gerarchia:** è na misura analoga alla precedente, che specifica quanto il vincolo complessivo si concentri su un solo Alter che, appunto, assume come tale un ruolo “gerarchico” nei confronti di Ego.



## MISURE DI RETE /15

### MISURE RELATIVE AD EGO

- **Brokerage (misure di Gould e Fernandez)**: Nel caso in cui si disponga di informazioni sugli attributi degli attori (ovvero sia possibile operare delle partizioni sulla matrice di adiacenza), è possibile compiere più sofisticate (e spesso “pericolose”) misure di brokerage, distinguendo 4 diversi ruoli di Ego, denominati rispettivamente:
  - **“coordinatore”**, quando Ego unisce due attori appartenenti alla sua medesima partizione (cioè dotati degli stessi attributi);
  - **“consulente”**, quando gli attori uniti da Ego hanno fra loro attributo uguale ma diverso da Ego;
  - **“gatekeeper”**, quando Ego intermedia la relazione “in ingresso” fra un attore ad esso eguale per attributo ed un attore dotato di altro attributo;
  - **“rappresentante”**, quando Ego intermedia la relazione “in uscita” fra un attore ed esso eguale per attributo ed un attore dotato di altro attributo;
  - **“legante”**, quando Ego intermedia la relazione fra due attori diverse per attributo tanto fra loro che verso Ego medesimo.



## MISURE DI RETE /16

### MISURE RELATIVE AI GRUPPI

**In generale**, un network può sempre essere visto come un insieme di networks più piccoli (che possiamo chiamare “gruppi”), fra loro connessi. L'identificazione dei gruppi è una fra le parti fondamentali della SNA, e consente di generare vettori di attributi di appartenenza dei singoli attori, a loro volta utili per approfondire le altre misure.

#### Approcci bottom/up

- **Cliques**: una clique è un insieme di attori fra loro connessi in modo più “stretto” di quanto lo sia in generale il network complessivo di cui essi fanno parte. Detto in termini più formali, una clique è il massimo numero di attori che hanno tutti i possibili legami che possono teoricamente esistere fra loro stessi (un subgrafo completo e massimo). Tale vincolo piuttosto rigido è di solito affrontato dal punto di vista del calcolo simmetrizzando preventivamente la matrice di adiacenza, riferendosi dunque ad un concetto di relazione “presente”, al di là del suo orientamento.
- Ogni clique viene espressa in termini di *i*) membri che la compongono; *ii*) proporzione di legami che ogni membro ha con le diverse cliques (valore che varia ovviamente da 1 – in caso di appartenenza ad una clique a 0, nel caso in cui non abbia alcuna relazione con alcun membro della clique in esame).
- Un altro modo di vedere il rapporto fra attori e cliques è costruire una matrice di adiacenza che mostri il numero di cliques comuni a due attori, maggiore il quale maggiore la “trasversalità” della relazione fra gli attori stessi;
- Un terzo modo, popolare in quanto intuitivo, è esprimere il rapporto fra attori attraverso una clusterizzazione gerarchica (vedi lettura “*How to Explain Hierarchical Clustering*”, fornita a parte), rappresentandola successivamente in forma grafica di albero o dendrogramma. Possono essere oggetto di clusterizzazione sia la matrice “attori/cliques” (che non è – ovviamente – una matrice di adiacenza), sia la matrice “attori/attori” indicante il numero di cliques ad essi comuni.

## MISURE DI RETE /17

### MISURE RELATIVE AI GRUPPI

- **N-cliques.** Sono una generalizzazione delle cliques, sostituendo al vincolo forte del subgrafo completo e massimo l'esistenza di una relazione fra tutti gli attori attraverso un percorso di lunghezza massima  $N$ . Usualmente in sociologia si usano  $N$ -cliques con  $N$  posto a 2, ottenendo oggetti del tipo “amici degli amici”. Visto che rilassiamo il vincolo della lunghezza del legame, può essere utile esplorare diverse ipotesi di cliques agendo (in aumento) sul vincolo del numero minimo di membri (minimum size, posto da Ucinet per default eguale a 3).
- **N-clans.** E' una restrizione di  $N$ -cliques, attraverso il vincolo che il diametro sia minore o eguale a  $N$ . Corregge un difetto di  $N$ -cliques, quello di formare dei gruppi spuri, ovvero composti “casualmente” da membri che – molto più coerentemente – appartengono ognuno a gruppi fra loro differenti.
- **K-plexes.** Un altro approccio rivolto a rendere meno forte il vincolo originario delle cliques è quello di accettare come membro del gruppo ogni nodo che abbia almeno  $n - k$  legami con gli altri nodi, dove  $n$  è il numero complessivo dei nodi che costituiscono il gruppo. Altrimenti detto, il nodo  $A$  fa parte di un 2-plex costituito dai nodi  $B$ ,  $C$  e  $D$  quando p.e. ha un legame sia con  $B$  che con  $C$ , ma non con  $D$ , essendo  $D$  a sua volta legato sia a  $B$  che a  $C$ .  $K$ -plexes genera ovviamente molti più gruppi di minor dimensione che i metodi precedenti.
- **K-cores.** E' costituito da tutti gli attori connessi con almeno  $K$  altri attori nell'ambito del medesimo gruppo o se si preferisce, è un subgrafo completo il cui degree sia eguale o maggiore di  $K$ . Nota: la routine Ucinet che calcola i gruppi definisce  $K$  in automatico, partendo dal valore minimo 3 e procedendo fino a che non viene trovato alcun gruppo.
- **F-groups.** E' una modalità di costruzione dei gruppi rivolta a reti con legami valorizzati. Definisce gruppi mutuamente esclusivi sulla base della condizione di transitività debole, relativa ad un opportuno valore di cut off scelto dal ricercatore. Altrimenti detto, dati 3 nodi  $A$ ,  $B$  e  $C$ , essi sono membri di un gruppo se i legami  $AB$  e  $BC$  sono entrambi più forti del legame  $AC$ , a sua volta maggiore del valore di cutoff.

## MISURE DI RETE /18

### MISURE RELATIVE AI GRUPPI

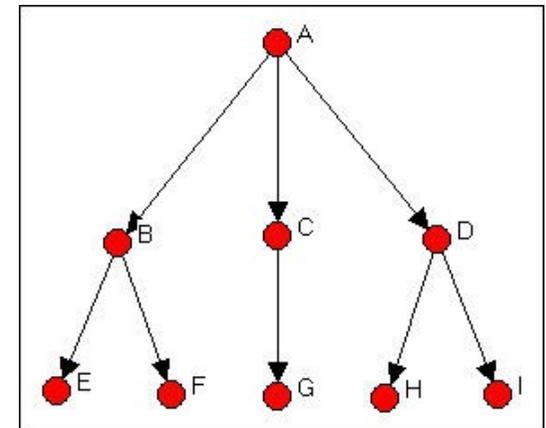
#### Approcci top/down

- **Analisi delle componenti.** Le componenti di un grafo sono subgrafi connessi al loro interno (oppure sono nodi isolati) ma non connessi fra di loro. Nel caso di reti binarie con legami orientati è possibile distinguere fra componenti “deboli” (quando i nodi sono connessi senza simmetria) e “forti” (il caso opposto). Nel caso di reti con legami valorizzati, le componenti sono determinate in modo iterativo, rilassando progressivamente il valore del cutoff (il grafo originale è dicotomizzato in  $n$  grafi, uno per ogni valore di cutoff, e per ognuno di essi sono determinate le componenti).
- **Blocchi e cutpoints.** Un altro approccio top/down è rivolto alla ricerca dei cutpoints, cioè dei nodi che – se rimossi – portano alla creazione di subgrafi disconnessi, detti “blocchi” o “bi-componenti”. Normalmente questa misura è preceduta – per i networks orientati – dalla simmetrizzazione dei legami. Dal punto di vista sociologico, questa misura esprime – nei suoi limiti – la vulnerabilità della rete.
- **Lambda sets.** In alternativa alla ricerca dei cutpoints, è possibile cercare i legami “ponte”, cioè quelli che – una volta interrotti – portano alla disconnessione del network in subgrafi. Una modalità più sofisticata di analisi è rivolta alla valutazione della riduzione del flusso massimo (vedi “misure di connessione”) che deriverebbe dallo scioglimento di ogni legame, in modo da determinare quelli maggiormente “importanti” rispetto alla capacità di comunicazione complessiva della rete. Anche in questo caso normalmente il network in esame è precedentemente simmetrizzato. Il risultato – denominato lambda set – esprime per ogni legame il grado di minima connettività interna, ed è rappresentato come cluster analysis gerarchica.
- **Fazioni.** Un'ultima misura di tipo top/down è data dalle “fazioni”, che esprimono la prossimità del network rispetto ad una forma idealtipica in cui: *i)* ogni attore appartiene ad un gruppo e, al contempo, *ii)* ogni gruppo è una componente (come a dire che la società ideale presa a riferimento è molto più basata sulla frammentazione che sull'integrazione).

## MISURE DI RETE /19

### MISURE RELATIVE A RUOLI E POSIZIONI

- **Equivalenza strutturale.** E' la forma di equivalenza più forte. Due attori sono equivalenti in senso strutturale se hanno esattamente gli stessi legami con tutti i nodi di cui si compone il network. Nel grafo a fianco, vi sono p.e. 7 classi di equivalenza strutturale, rispettivamente popolate dai nodi {A}; {B}; {C}; {D}; {G}; {E, F}; {H, I}. Sociologicamente, esprime il concetto “forte” di posizione, stante l'identità degli attributi strutturali.
- **Equivalenza automorfica.** E' una forma più rilassata di equivalenza, che richiede agli attori di essere posizionati nel network in maniera analoga (ovvero avere le medesime distanze dai medesimi tipi di attori; avere lo stesso pattern di legami), anche in assenza di relazioni effettivamente comuni. Nel grafo a fianco, vi sono p.e. 5 classi di equivalenza automorfica, rispettivamente popolate dai nodi {A}; {C}; {B, D}; {G}; {E, F, H, I}. Sociologicamente, esprime il concetto “generale” di posizione, cioè di rapporto “spaziale” fra attori.
- **Equivalenza regolare.** E' una forma ancora più rilassata di equivalenza (e dunque maggiormente utile, visto che nella SNA spesso la semplificazione porta ad una maggiore comprensione delle reti), che richiede agli attori di avere lo stesso tipo di legami con altri attori, a loro volta dotati delle caratteristiche di equivalenza regolare. Nel grafo a fianco, vi sono p.e. 3 classi di equivalenza regolare, rispettivamente popolate dai nodi {A}; {B, C, D}; {E, F, G, H, I}. Sociologicamente, esprime il concetto di ruolo, cioè di rapporto fra tipologie (classi) di attori.



## MISURE DI RETE /20

### MISURE RELATIVE A RUOLI E POSIZIONI

Nei networks reali è raro che si ritrovino con grande frequenza casi “assoluti” di equivalenza; è invece più probabile che interessino misure di “grado di equivalenza”, cioè di “prossimità” alle equivalenze teoriche.

#### Misure generali di similarità/dissimilarità

- In primo luogo, va costruita una matrice di adiacenza di righe e colonne concatenate, aggiungendo alle righe originarie tutte le colonne copiate come righe. Le misure di similarità fra due attori sono così ridotte al confronto fra i due vettori che li esprimono in modo ora completo. Naturalmente è sempre possibile compiere misure riferite solo ai valori di riga o di colonna. Vi sono diverse modalità di confronto:
  - **Coefficiente di relazione di Pearson**. Esprime il grado associazione lineare fra due variabili. Varia fra -1 (due attori hanno legami opposti verso gli attori cui sono connessi), 0 (dalla conoscenza del legame di un attore non posso probabilisticamente risalire a quella dell'altro attore, ovvero non vi è correlazione) e +1 (identità). Pearson va di solito applicato a legami valorizzati di tipo scalare o ordinale;
  - **Distanza euclidea**, ovvero radice quadrata della somma delle differenze elevate al quadrato. Altrimenti detto: l'intensità del legame dell'attore A con C è sottratta dall'intensità del legame dell'attore B con C, e la differenza è elevata al quadrato. Ciò è ripetuto per tutti gli attori, i risultati sono sommati ed è estratta la radice quadrata del risultato. Le distanze euclidee sono utilizzabili sia per legami binari che valorizzati.
  - **Match esatto**: tipica misura di reti binarie. Esprime il numero di volte in cui il legame che un attore A ha (o non ha) con gli altri attori è lo stesso che B ha con gli stessi attori, normalizzato rispetto al numero dei legami possibili.

## MISURE DI RETE /21

### MISURE RELATIVE A RUOLI E POSIZIONI

- **Misura di Jaccard**: esprime il numero di volte che A e B hanno un legame con lo stesso attore terzo, come percentuale del numero totale dei legami ad esso diretti. Rispetto alla misura precedente si ignorano i casi in cui sia A che B non sono legati al medesimo nodo C e si chiede, per il totale dei legami presenti, la percentuale in comune.
- **Distanza di Hamming**: è il numero di legami che devono essere cambiati nel vettore di un nodo per renderlo identico al vettore di un altro nodo.

### Rappresentazioni attraverso *cluster analysis*

- Normalmente le misure di similarità/dissimilarità sono rappresentate attraverso cluster analysis. Abbiamo già visto in apposita lettura che i cluster sono formati attraverso un processo iterativo basato sulla progressiva riduzione delle distanze fra i membri. Diversi sono i modi che possono essere utilizzati per misurare tali distanze:
  - **legame singolo** (o metodo della connettività, o “vicino più prossimo”). La distanza fra due cluster è definita come la minor dissimilarità (la maggior similarità) fra i membri. Altrimenti detto, nel metodo a legame singolo le similarità sono computate sulla base della similarità del membro del nuovo cluster che è maggiormente simile ad ogni altro caso non compreso nel cluster;
  - **legame completo** (o metodo del diametro, o “vicino più lontano”). La distanza fra due cluster è definita come la maggior dissimilarità (la minor similarità) fra i membri. Altrimenti detto, nel metodo a legame completo le similarità sono computate sulla base della similarità del membro del nuovo cluster che è il meno simile ad ogni altro caso non compreso nel cluster;
  - **media**: la distanza fra due cluster è definita come la dissimilarità (o similarità) media fra i membri.

## MISURE DI RETE /22

### MISURE RELATIVE A RUOLI E POSIZIONI

- Il valore indicato alla voce “livello” esprime il grado di associazione (similarità o dissimilarità) fra gli item che costituiscono il cluster. Se utilizziamo come metodo di clusterizzazione il “legame semplice” e applichiamo l'analisi su una matrice di similarità, il livello con valore  $x$  significa che qualunque item contenuto nel cluster è almeno simile  $x$  unità rispetto ad almeno un altro item costituente il medesimo cluster. Se si utilizza il metodo “legame completo” significa che è simile  $x$  unità rispetto a qualunque altro componente il cluster. Se, infine, si usa il metodo della “media”  $x$  esprime appunto la simiglianza media fra i membri del cluster.

### Rappresentazioni multidimensional scaling (MDS)

- Una matrice di prossimità, ma anche una normale matrice di adiacenza, possono essere rappresentate come punti (nodi) proiettati in uno spazio a  $k$  dimensioni, tale per cui le distanze euclidee fra i punti corrisponda il più possibile all'ordinamento delle prossimità. Altrimenti detto, MDS rappresenta il pattern di similarità/dissimilarità nel profilo dei legami fra gli attori come un diagramma di dispersione in uno spazio multidimensionale, in modo da rendere visivamente quanto ogni attore sia prossimo agli altri. Le dimensioni costruite dai diversi algoritmi applicabili, per quanto abbiano significato matematico, richiedono di essere interpretate in senso sociologico, non necessariamente avendo “senso”. Inoltre, è bene sapere che gli algoritmi non sempre portano alle stesse soluzioni (possono esistere più minimi locali fra loro equivalenti in senso numerico, anche se di significato qualitativo differente) il che richiede di utilizzare tentativamente l'algoritmo, e valutare comparativamente i risultati. Un indicatore importante (anche se non esaustivo) è il grado dello stress (cioè della fitness dei valori), di cui ci si attende normalmente valori inferiori a 0,15.

## MISURE DI RETE /23

### MISURE RELATIVE A RUOLI E POSIZIONI

#### Misure di equivalenza strutturale

- **Misura del profilo di similarità/distanza.** Come visto, due attori sono strutturalmente equivalenti se hanno lo stesso pattern di legami con gli altri attori cui sono connessi. Il che equivale a confrontare, coppia a coppia, i vettori che descrivono i legami in ingresso ed in uscita da ogni attore e esprimere la loro similarità o – al contrario – la loro distanza. Il vettore del profilo è definito come la concatenazione di ogni profilo del vettore delle relazioni individuali. Se si lavora su una matrice di adiacenza originaria (cioè costituita dai legami fra gli attori), la similarità può essere computata sul profilo dei legami (p.e. attraverso le misure “match esatto” e “Coefficiente di Jaccard”). Se invece si trasforma preventivamente la matrice di adiacenza in una matrice di (dis)similarità (p.e. calcolando la distanza geodetica fra gli attori) si possono utilizzare misure quali il coefficiente di relazione di Pearson o la distanza euclidea.
- **CONCOR:** è un altro algoritmo usato per compiere misure di equivalenza strutturale. Esso esamina coppia per coppia, quanto sono simili i vettori di similarità di tutti gli attori, costruendo una matrice di correlazione basata sul coefficiente di Pearson. Su tale matrice è applicata iterativamente la stessa procedura, fino alla convergenza. Il ricercatore definisce ad inizio analisi il livello di *splitting* fra i blocchi, ovvero quanto è “profondo” l'albero delle cluster. L'obiettivo di CONCOR è mettere in evidenza i principali gruppi di attori fra loro sufficientemente equivalenti, in modo da “partire” la matrice in blocchi ed esaminare che tipi di relazioni si istituiscono fra di essi.

## MISURE DI RETE /23

### MISURE RELATIVE A RUOLI E POSIZIONI

- Un altro approccio simile è l'ottimizzazione via **Tabu search**. Il network è inizialmente diviso in modo arbitrario nel numero di blocchi definito per ipotesi dal ricercatore. Se i blocchi rappresentassero le diverse classi di equivalenza strutturale, ognuno di essi conterrebbe solo i valori 0 o 1. La distanza da tale situazione ideale (detta “numero di errori”) corrisponde al minimo numero di cambiamenti che andrebbero fatti per riportare la realtà all'idealtipo di riferimento. La somma del numero degli errori misura la funzione di costo del grado di equivalenza strutturale. L'algoritmo Tabu opera in modo iterativo, cercando il minimo della funzione di costo, ovvero la miglior partizione di equivalenza strutturale. I blocchi così ottenuti si caratterizzano ovviamente per avere la varianza minima.

#### Misure di equivalenza automorfica

- Invece di chiedersi – come nel caso dell'equivalenza regolare – quali individui possono essere scambiati nel network senza modificare le relazioni sociali, l'equivalenza automorfica si domanda quali subgrafi (cioè gruppi di individui vincolati fra loro da relazioni) possono essere “scambiati” con altri subgrafi, mantenendo invariate le proprietà complessive della rete. Ovvero quante posizioni equivalenti esistono nel network. Come sempre, il vero obiettivo non è trovare le equivalenze esatte, quanto utilizzare il concetto di equivalenza per costruire delle classi mutuamente escludenti di attori fra loro sufficientemente simili e – al contempo – il più possibile distanti dalle altre classi.
- La prima e più ovvia strategia di ricerca è comparare fra loro tutte le combinazioni di possibili **permutazioni** del network, misurando per ognuna di esse quanto la distanza fra ogni paio di nodi del nuovo (sub)grafo è prossima a quella del (sub)grafo originario. Quando il nuovo (sub)grafo ha le stesse distanze di quell'originario, ovvero può essere sovrapposto esattamente ad esso, significa che è ad esso isomorfo. Le diverse classi di isomorfismo sono denominate “orbite”, e la loro ricerca può richiedere – in caso di reti di grandi dimensioni – tempi non proprio ragionevoli.

## MISURE DI RETE /24

### MISURE RELATIVE A RUOLI E POSIZIONI

- Un secondo approccio, maggiormente veloce ma meno preciso è basato sull'algoritmo **Tabu** che abbiamo già incontrato. Il ricercatore fissa ipoteticamente il numero delle classi di isomorfismo attese e l'algoritmo inizia ad allocare in modo casuale i nodi nelle partizioni. Successivamente è misurata la bontà della fitness, come somma della varianza dei singoli blocchi, questa data dalla somma dei quadrati per ogni riga e colonna contenute in ogni blocco. L'algoritmo procede in modo iterativo a modificare la composizione dei singoli blocchi, fino a minimizzare la varianza complessiva. I blocchi risultanti sono caratterizzati dalla medesima varianza interna, cioè i nodi in essi contenuti hanno una elevata probabilità di avere il medesimo profilo di legami in ingresso ed in uscita e di avere legami di eguale stessa distanza dagli attori contenuti nelle altre classi. Come sempre nel caso di ricorso ad algoritmi che cercano tentativamente di minimizzare/massimizzare funzioni, è utile eseguire i calcoli più volte, in modo da verificare la solidità del risultato proposto.
- Nel caso in cui si disponga di network con legami valorizzati, è possibile generalizzare il concetto di automorfismo attraverso opportune approssimazioni. L'algoritmo **Maxsim** inizia a calcolare il reciproco della distanza fra gli attori, assumendo il valore del legame come “forza” della relazione. Successivamente, le distanze di ogni attore sono riorganizzate in una lista ordinata degressivamente. Per ogni coppia i attori è costruito un profilo di dissimilarità, attraverso misura della distanza euclidea. Gli attori fra loro più prossimi sono considerati fra loro maggiormente automorfici. Infine, una cluster analysis è svolta per definire le classi di automorfismo “approssimato” così rilevate. L'algoritmo di massima similarità è applicabile anche a reti con legami binari, previa conversione degli stessi in distanze geodetiche fra i nodi.

## MISURE DI RETE /24

### MISURE RELATIVE A RUOLI E POSIZIONI

#### Misure di equivalenza regolare

- Come si è visto, l'equivalenza regolare – a cui in sociologia corrisponde il concetto di ruolo – è la condizione più rilassata di equivalenza, ed è definita secondo una logica di “interdipendenza” fra gli attori: due attori presentano infatti equivalenza regolare se sono egualmente relazionati ad altri attori fra loro dotati di equivalenza regolare. Il che, dal punto di vista pratico, significa osservare i legami in ingresso ed in uscita dai diversi attori, definendo un vettore di partizione della matrice di adiacenza che la riorganizzi in gruppi, ognuno dei quali contenente attori eguali dal punto di vista del tipo di rapporto con gli altri gruppi. Diversi algoritmi provvedono a partizionare la matrice in blocchi rispondenti a tale criterio.
- **Categorical REGE.** E' un algoritmo iterativo rivolto a matrici costituite da legami binari o espressi come variabili nominali. Nella prima iterazione, CATREGE classifica i nodi in tre tipologie, sulla base del tipo di legami che li caratterizzano (sorgenti, ripetitori, riceventi). Nella seconda iterazione ogni nodo è classificato sulla base della tipologia dei suoi vicini: se due nodi hanno tipologie di vicini diversi, sono classificati in due partizioni diverse. Ogni successiva iterazione utilizza come riferimento il risultato della classificazione dello step precedente. Il procedimento ha termine quando tutti i nodi sono attribuiti ad una categoria, il numero massimo delle quali è pari al numero dei nodi di cui il network si compone. L'esito del procedimento è una matrice di similarità che mostra a quale livello del processo di iterazione ogni nodo si è “separato” dagli altri, costituendo una categoria. Come sempre, la matrice è più utilmente rappresentabile attraverso una cluster analysis.

## MISURE DI RETE /25

### MISURE RELATIVE A RUOLI E POSIZIONI

- **REGE**. E' la versione di Categorical REGE per legami valorizzati nel continuo. Per ogni coppia di nodi, la funzione di confronto è basata sull'esame della similarità dei rispettivi vicini (misurata sulla base del valore assoluto della differenza della forza dei legami), attraverso l'usuale procedimento iterativo di cui è fissato a priori il numero di passi.
- **Nucleo/Periferia**. Un'ultima misura che non appartiene strettamente al concetto di equivalenza ma non è distante da un'idea di “ruolo” risponde all'idea di dividere nodi di un network in due tipologie fisse:
  - il “nucleo”, cui appartengono gli attori che hanno una elevata densità di relazioni fra loro stessi;
  - la “periferia”, ovvero gli attori che hanno scarsi legami fra loro medesimi, ovvero sono ai margini della rete.

Sono disponibili algoritmi genetici di partizione Nucleo/Periferia sia per legami binari che valorizzati, come per reti mono o bimodali. La fitness è espressa in una scala 0-1, dove (differentemente dagli algoritmi iterativi prima visti) il valore 1 corrisponde al best fit.

## BIBLIOGRAFIA ESSENZIALE

### TESTI DI RIFERIMENTO

Carrington P., Scott J., Wasserman S. (2005), *Models and Methods in Social Network Analysis*, Cambridge (MA), Cambridge University Press.

Salvini, A. (2007), *Analisi delle reti sociali. Teorie, metodi, applicazioni*, Milano, Franco Angeli.

Wasserman S., Faust K., (1996), *Social Network Analysis. Method and Applications*, Cambridge (MA), Cambridge University Press.

### RIVISTE DI RIFERIMENTO

Connections

Social Networks

### ALCUNI SITI DI RIFERIMENTO

- <http://www.insna.org/>
- <http://www.socialnetworkanalysis.com/>