



**UNIVERSITÀ DEGLI STUDI DI BARI “ALDO MORO”**

DIPARTIMENTO INTERATENEO DI FISICA

CORSO DI LAUREA DI I LIVELLO IN FISICA

---

**RETI COMPLESSE E CONNETTIVITÀ  
STRUTTURALE DEL CERVELLO  
UMANO**

**Relatore:**

Chiar.mo Prof. Sebastiano STRAMAGLIA

**Laureando:**

Vito Antonio Rocco SUSCA

---

ANNO ACCADEMICO 2012-2013

*Ai miei genitori*

# Indice

|                                                                                                       |           |
|-------------------------------------------------------------------------------------------------------|-----------|
| <b>Introduzione</b>                                                                                   | <b>3</b>  |
| <b>1 Teoria dei grafi</b>                                                                             | <b>8</b>  |
| 1.1 Definizioni e notazioni . . . . .                                                                 | 8         |
| 1.1.1 Grafi, nodi e archi . . . . .                                                                   | 8         |
| 1.1.2 Tipi di grafi: orientati e non orientati, pesati e non pesati . . .                             | 8         |
| 1.1.3 Proprietà dei nodi: matrice di adiacenza, grado di un nodo e assortatività . . . . .            | 10        |
| 1.1.4 Proprietà dei collegamenti tra nodi: walk, path, distanza, connessione di un grafo . . . . .    | 14        |
| 1.2 Misure caratteristiche delle reti . . . . .                                                       | 15        |
| 1.2.1 Misure di segregazione locale: clustering, motivi e modularità                                  | 16        |
| 1.2.2 Misure di integrazione globale: cammino caratteristico ed efficienza globale . . . . .          | 23        |
| 1.2.3 Misure di centralità dei nodi: closeness e betweenness, hubs .                                  | 25        |
| <b>2 Topologia delle reti e modelli</b>                                                               | <b>28</b> |
| 2.1 Modelli ed architetture di reti ideali: reti random e reti regolari . . .                         | 28        |
| 2.1.1 Grafi random o grafi ER . . . . .                                                               | 28        |
| 2.1.2 Grafi regolari . . . . .                                                                        | 32        |
| 2.2 Proprietà emergenti delle reti reali: proprietà di piccolo mondo ed invarianza di scala . . . . . | 32        |
| 2.2.1 Effetto di “piccolo mondo” . . . . .                                                            | 33        |
| 2.2.2 Invarianza di scala: distribuzioni a legge di potenza . . . . .                                 | 34        |
| 2.3 Modelli ed architetture delle reti reali: modello WS e modello BA . .                             | 36        |
| 2.3.1 Un modello di rete di piccolo mondo: il modello di Watts e Strogatz . . . . .                   | 36        |
| 2.3.2 Un modello di rete ad invarianza di scala: il modello di Barabasi e Albert . . . . .            | 38        |
| 2.3.3 Il problema della modularità e la classificazione delle reti . . .                              | 39        |

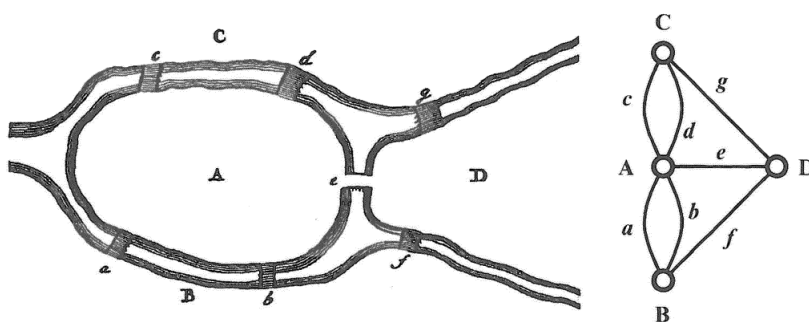
|          |                                                                                      |           |
|----------|--------------------------------------------------------------------------------------|-----------|
| <b>3</b> | <b>Connettività del cervello umano: il connettoma</b>                                | <b>43</b> |
| 3.1      | La prospettiva delle reti complesse nello studio del cervello umano . .              | 43        |
| 3.1.1    | Motivazioni . . . . .                                                                | 43        |
| 3.1.2    | Tipi di connettività . . . . .                                                       | 44        |
| 3.2      | Il connettoma umano . . . . .                                                        | 46        |
| 3.2.1    | Definizione e caratteri fondamentali. L'esempio del C. Elegans                       | 46        |
| 3.2.2    | Un esempio di connettoma: la mappa neurale del C. Elegans .                          | 48        |
| 3.3      | Tecniche sperimentali . . . . .                                                      | 49        |
| 3.3.1    | La dimensionalità multiscala del cervello umano . . . . .                            | 50        |
| 3.3.2    | Tecniche di imaging per la connettività su larga scala . . . . .                     | 52        |
| 3.3.3    | Procedura di inferenza del network strutturale dai dati . . . .                      | 55        |
| <b>4</b> | <b>Architettura delle reti cerebrali strutturali della corteccia cerebrale umana</b> | <b>59</b> |
| 4.1      | I principali studi sulla connettività cerebrale umana a larga scala . . .            | 59        |
| 4.2      | Le reti cerebrali strutturali non sono random . . . . .                              | 61        |
| 4.2.1    | Proprietà di piccolo mondo . . . . .                                                 | 61        |
| 4.2.2    | Modularità e hubs . . . . .                                                          | 67        |
|          | <b>Conclusioni</b>                                                                   | <b>71</b> |
|          | <b>Riferimenti bibliografici</b>                                                     | <b>73</b> |

# Introduzione

Le realtà è organizzata secondo intricati schemi di interconnessioni, a vari livelli: l'uomo è un essere sociale e l'esistenza di legami e relazioni, in generale connessioni tra individui, è alla base di qualunque società che garantisca la sopravvivenza agli membri che la compongono. Ciascun individuo è quindi punto nodale di una rete di rapporti e influenze, è parte di un complesso schema di connessioni di vario genere ed il suo peso all'interno di un tale pattern è tanto maggiore quanto più riesce a generare, in concerto con altri, moti collettivi o ad intercettare flussi preesistenti. Le reti pervadono l'esistenza di un individuo sottoforma di intricate connessioni tra persone, oggetti, informazioni. Non solo le relazioni sociali di base come amicizie e collaborazioni professionali, ma anche per esempio le svariate modalità di comunicazione, come la rete telefonica fissa, mobile, internet, la rete elettrica, il sistema dei trasporti (strade, rete ferroviaria, rotte aeree, il sistema degli aeroporti e stazioni), il genoma umano, gli schemi di sintesi delle proteine, ed il sistema nervoso sono tutti esempi di sistemi complessi cioè sistemi formati da un grande numero di componenti interagenti fra loro: la struttura irregolare, il comportamento e l'evoluzione nel tempo di questo genere di sistemi deriva dalle modalità di interazione tra le componenti e non solo dalle proprietà delle unità elementari in cui sono scomponibili (e.g. un'autostrada non ha una importanza definita di per sé stessa: le sue caratteristiche invece dipendono da come essa è collegata al resto della rete stradale e da quanto traffico è capace di sostenere). Risulta evidente quindi che lo studio di questi sistemi tramite oggetti matematici che, prescindendo dalla loro natura specifica e dalle dimensioni spaziali, ne analizzino la natura delle connessioni tra le varie parti che li compongono dal punto di vista topologico sia un approccio naturale e quasi automatico. Tali realtà possono essere formalizzate in termini di reti complesse e studiate dalla scienza delle reti, una disciplina a cavallo tra la matematica (linguaggio e formalismo della teoria di grafi), la fisica (attitudine al problem solving attraverso analogie, test statistici e tecniche di indagine) e l'informatica (gestione di database e potere computazionale), che negli ultimi decenni ha acquisito sempre maggior importanza.

Il tasso crescente di interesse verso lo studio dei networks complessi negli ultimi

anni è dovuto essenzialmente alla disponibilità di numerosi e vasti database da cui attingere dati e risorse e allo sviluppo delle tecnologie informatiche, che forniscono mezzi di calcolo sempre più potenti necessari ad affrontare uno studio di questo tipo dato l'enorme numero di elementi interagenti (Un esempio: il world wide web è la rete formata dagli hyperlinks tra le diverse pagine web e con più di  $10^8$  nodi, è la rete più grande mai studiata) [1]. L'approccio "filosofico" al problema delle reti è comunque ben radicato nella storia. Il primo ad utilizzare uno schema interpretativo simile per risolvere in primis un problema pratico e poi formulando un intero linguaggio atto a descrivere un'intera classe di problemi geometrici fu L. Eulero nel 1736, affrontando il cosiddetto "problema dei ponti di Konisberg" (Fig. 1) [2]. In breve, Eulero riuscì in termini esatti a dirimere il rompicapo che chiedeva se esistesse un percorso che cominciava e terminava nello stesso punto tale che ciascuno dei 7 ponti della città di Konisberg venisse stato attraversato una sola volta. Era opinione comune che tale percorso non esistesse, ma Eulero, formalizzando la questione in termini di "geometria situs" ossia geometria di posizione, dimostrò con esattezza che effettivamente un tale percorso non poteva esistere e pertanto che la non esistenza del percorso non era legata all'incapacità di trovarlo ma era una precisa proprietà del grafo. La geometria di posizione di cui parla il matematico svizzero, una geometria che non necessita di "misure di distanze né calcoli", ma solo di "determinazioni di posizione e delle sue proprietà" altro non è che il prototipo della teoria dei grafi, cioè l'ossatura matematica della teoria delle reti complesse.



**Figura 1:** Ponti di Konisberg

Nei secoli successivi, sino ai primi anni del '900, la teoria dei grafi si è sviluppata nell'ambito della matematica discreta e sono state per prime le scienze sociali ad appropriarsi del linguaggio dei grafi per analizzare le connessioni e le modalità di comunicazione tra gruppi di individui a varie scale: da piccoli gruppi coesi come per esempio i dipendenti di una stessa azienda, sino all'analisi dei commerci tra nazioni. Il celebre studio di Milgram sui gradi di separazione tra due cittadini scelti random residenti in stati diversi degli Usa si colloca in questo contesto. L'esperimento, basato su una catena con lo scopo di far recapitare una lettera ad un destinatario non conosciuto direttamente dal mittente tramite passaggi intermedi tra individui che si conoscevano reciprocamente, fa sorridere se calato nella realtà attuale, basata sulle comunicazioni via Internet e sui social networks, una tra le evidenze più schiaccianti del fatto che le reti sono un fatto costituente della nostra vita.

In ambito prettamente scientifico, le prime manifestazioni della teoria dei grafi come strumento interpretativo di realtà complesse è rintracciabile nel lavoro di Kirchhoff che elaborò i celebri teoremi nei nodi e delle maglie a partire dalla rappresentazione delle reti elettriche tramite grafi. Analogamente in chimica, il matematico Cayley fu il primo a descrivere gli isomeri degli idrocarburi, molecole con lo stesso numero di atomi che differiscono per il modo in cui questi atomi sono legati fra di loro, tramite l'uso di "grafi chimici". Le applicazioni moderne della teoria dei grafi per lo studio delle reti sono svariate e possono applicarsi a qualsiasi sistema complesso come cellule, cervello, ecosistemi, società ed economia globale. L'approccio più recente, la cosiddetta "nuova scienza delle reti" consiste nel caratterizzare e descrivere la struttura dei sistemi complessi, determinata dalla connettività fra i loro elementi e fare delle predizioni teoriche circa la loro evoluzione e funzionalità. La potenza del metodo sta proprio nella sua generalità e versatilità di applicazione alle casistiche più disparate ed alla capacità di sintetizzare in poche misure topologiche fondamentali caratteristiche generali che non potrebbero emergere attraverso un'analisi di tipo riduzionistico di mera scomposizione e studio del sistema nelle sue singole entità. In altre parole, molti sistemi complessi mostrano comportamenti macroscopici molto simili fra loro nonostante le profonde differenze tra gli elementi microscopici che li costituiscono e i modi con cui essi interagiscono: il comportamento dei più disparati

sistemi complessi infatti è plasmato dal modo in cui si interconnettono gli elementi costitutivi e tali modalità sono generali cioè non appartengono alla specifica tipologia di network.

L'idea di studiare il cervello e in generale il sistema nervoso tramite le reti complesse deriva da osservazioni di carattere anatomico e funzionale, emerse sin dal diciannovesimo secolo [3]. Gli studi sistematici dell'anatomia del sistema nervoso a partire da organismi semplici come i nematodi sino a giungere ai mammiferi e quindi all'uomo hanno mostrato che esso ha una struttura di rete complessa a diverse scale. A livello cellulare, il sistema nervoso può essere analizzato come la rete creata dal vasto numero di unità base, i neuroni, connessi fra loro tramite le sinapsi. Ad una scala intermedia, può essere rappresentato come l'interazione fra gruppi di neuroni che formano aree specializzate in una funzione specifica. Ad una scala più grande, si può considerare il sistema nervoso come la rete composta dalle varie aree funzionali del cervello che comunicano fra loro attraverso i fasci di fibre nervose (sostanza bianca), che costituiscono il substrato anatomico delle vie neuronali di tipo funzionale. L'idea di una serie di elementi costituenti connessi fra loro quindi è lo schema di base dei sistemi neurali che si ripete a diverse scale: scaturisce in modo naturale l'utilizzo dello strumento delle reti cerebrali come strumento interpretativo di un sistema così organizzato. Un'altra evidenza che porta all'utilizzo delle reti per lo studio del cervello è data da osservazioni di tipo funzionale. Attività non elementari come l'elaborazione di risposte agli stimoli, l'attuazione di strategie, lo svolgimento di compiti, la memoria e l'auto-rappresentazione non sono a carico di una singola cellula nervosa ma possono avvenire solo nell'ambito di una cooperazione tra un largo numero di neuroni connessi in modo opportuno. Ciò mette in luce la funzionalità integrativa del cervello ossia il fatto che la sua fisiologia è legata strettamente al modo in cui le sue componenti interagiscono.

Obiettivo di questo lavoro è la descrizione fondamentale della cinematica delle reti cerebrali, ossia l'analisi delle proprietà topologiche salienti della connettività *strutturale* del cervello umano, tramite le tecniche di misura (da non intendersi in senso metrico) delle reti complesse. L'analisi della connettività *funzionale* ed *efficace*, ossia la dinamica e l'evoluzione temporale delle reti cerebrali, non sarà trattata: le

due direzioni di ricerca sono fortemente legate e procedono parallelamente ma si è preferito evidenziare l'importanza dell'aspetto strutturale ossia la realizzazione e lo studio del *Connettoma* umano, una mappa completa delle connessioni anatomiche del cervello umano, come fondamento meccanicistico e strumento di comprensione di ogni pattern dinamico sovrastante. Il lavoro di tesi sarà organizzato secondo uno schema bipartito: nei capp. 1 e 2 verrà esposta la teoria essenziale delle reti complesse inquadrata nella cornice matematica della teoria dei grafi. Nel cap. 1 saranno date le definizioni e le notazioni utili a caratterizzare i componenti di base dei grafi, le loro proprietà caratteristiche e le tipiche misure che ne specificano il carattere. Nel cap. 2 verranno spiegati e classificati i modelli principali delle reti complesse, partendo da modelli prettamente matematici sino ad arrivare modelli più aderenti alla realtà e adeguati alla descrizione di reti reali.

Nel cap. 3, sarà affrontata la “cinematica” dei brain networks: si definiranno i tipi di connettività riscontrabili nel cervello umano, verranno passate in rassegna le metodologie di imaging e le procedure sperimentali grazie alle quali è possibile estrapolare dall'encefalo il suo network strutturale complessivo, il connettoma, mettendo in evidenza le problematiche di questo tipo di studio e operando un confronto con l'unico esempio di connettoma completamente conosciuto, quello del nematode *Caenorhabditis elegans*. Nel cap. 4, infine, saranno esposte le proprietà salienti della connettività strutturale umana, che emergono dalla sua analisi effettuata mediante l'utilizzo delle tecniche di “misura” topologica delle reti complesse esposte nei capitoli precedenti.

# 1 Teoria dei grafi

## 1.1 Definizioni e notazioni

La trattazione matematica delle reti complesse avviene nel ramo della matematica discreta denominato teoria dei grafi: formalmente, un grafo è la rappresentazione in termini matematici di una rete complessa del mondo reale. Posta questa differenza tra modello matematico e oggetto reale modellizzato, i due termini sono confondibili.

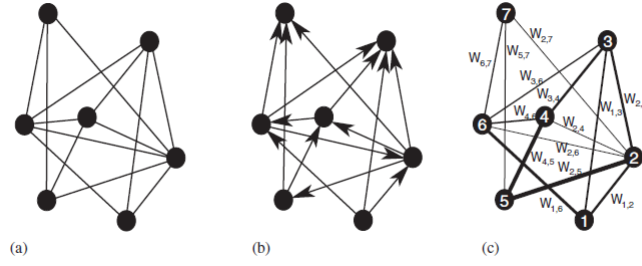
### 1.1.1 Grafi, nodi e archi

Un grafo  $G = (N, L)$  consiste in una coppia di due insiemi tali che  $N \neq \emptyset$  e  $L$  è un insieme di coppie di elementi di  $N$  [1]. Gli elementi di  $N = \{n_1, n_2, \dots, n_N\}$  sono detti *nodi* (o *vertici*) del grafo  $G$  e rappresentano le unità fondamentali del sistema, mentre gli elementi di  $L = \{l_1, l_2, \dots, l_K\}$  sono detti *archi* o *links* e rappresentano le connessioni tra coppie di nodi ossia le relazioni tra gli elementi del sistema.  $N$  è il numero degli elementi di  $N$  ossia il numero di nodi ed indica la *dimensione* del grafo mentre  $K$  è il numero degli elementi di  $L$  ossia il numero di archi: il grafo  $G$  si potrà indicare come  $G = (N, L) = G(N, K) = G_{N,K}$  indifferentemente. Un nodo è indicato con la sua posizione  $i$ -esima nell'insieme  $N$ , mentre un arco, definito dalla coppia di nodi  $i$  e  $j$  che connette, si indica come  $l_{ij} = (i, j)$ : si dice anche che il link  $l_{ij}$  è incidente nei nodi  $i$  e  $j$ . Due nodi connessi da un link sono detti *adiacenti* o *vicini*.

### 1.1.2 Tipi di grafi: orientati e non orientati, pesati e non pesati

A seconda del tipo di archi, si possono individuare diversi tipi di grafo [1]. Gli archi si possono distinguere a seconda della direzione in cui connettono due nodi.

1. Un arco è detto *orientato* se gli elementi di  $L$  ossia coppie di nodi sono coppie ordinate di elementi di  $N$ : in questo caso,  $l_{ij}$  indica un arco diretto da  $i$  a  $j$  e  $l_{ij} \neq l_{ji}$ .



**Figura 2:** a) grafo non orientato; b) grafo orientato; c) grafo pesato

2. Un arco è detto *non orientato* se gli elementi di  $L$  sono coppie non ordinate di elementi di  $N$ : in questo caso, non esiste un'orientazione privilegiata e  $l_{ij} = l_{ji}$ .

D'ora in poi, in questa trattazione, il termine grafo privo di aggettivi indicherà un grafo non orientato. Per un grafo (non orientato)  $G$  di dimensione  $N$ , il numero di archi  $K$  può variare tra:

- $K = 0$ : valore minimo, quando nessun nodo è connesso a nessun altro;
- $K = N(N-1)/2$ : valore massimo, quando tutti i nodi sono adiacenti a coppie.

Infatti, dati  $N$  nodi e scelto a caso un nodo, si possono formare  $N-1$  coppie: l'operazione si può ripetere per ciascuno degli  $N$  nodi scelti a caso come primo membro della coppia. Dato che le coppie simmetriche sono identiche, si introduce il fattore  $\frac{1}{2}$  per eliminare le ridondanze, giungendo al risultato sopra indicato. Analogamente, si sarebbe potuto procedere considerando il coefficiente binomiale  $K = \binom{N}{2} = \frac{N!}{2!(N-2)!}$ . Un grafo è detto *rado* se  $K \ll N^2$ , mentre è detto *denso* se  $K = O(N^2)$ : se  $K = \frac{N(N-1)}{2}$ ,  $G$  è detto  $N$ -grafo completo ed si denota con  $K_N$ . Da un punto di vista grafico, un grafo può essere rappresentato da un insieme di punti ognuno indicante ciascun nodo, connessi da tratti o frecce nel caso di un grafo orientato. Ad esempio,  $K_3$  è un triangolo.

Inoltre, gli archi possono essere distinti in due categorie, a seconda che si considerino o meno l'intensità della connessione tra due nodi. Un arco può essere:

1. *non pesato* o *binario*: un arco che connette due nodi può essere presente o non presente, come un interruttore per il quale sono possibili solo le due posizioni

di on (1) ed off (0). Di conseguenza, non esiste una scala di importanza tra i vari links ma, quando presenti, sono tutti equivalenti fra loro se classificati per la loro importanza nel network.

2. *pesato*: a ciascun arco è attribuito un numero reale  $w_{ij}$  che determina il suo peso ossia la forza della connessione tra due nodi, rendendo così possibile la distinzione tra legami deboli e legami più forti. Ciò è utile in particolari tipi di reti reali come quelle sociali, le reti di trasmissione di segnale elettrica, il World Wide Web e le stesse reti cerebrali, in generale in tutti casi dove la pura descrizione topologica non rende giustizia di tutte le proprietà osservate, benchè tutta la teoria classica dei grafi sia stata principalmente sviluppata per grafi non orientati di tipo binario [1,3].

La teoria dei grafi binari e quella dei grafi pesati non differiscono di molto se non nella necessità di integrare il parametro del “peso” a partire dalle definizioni di base e in alcune misure topologiche fondamentali, generalizzandole: la definizione di grafo data in precedenza è infatti relativa ad un grafo non pesato.

Si definisce *grafo pesato* un ensemble di tre insiemi  $G=(N, L, W)$  tali che  $N \neq \emptyset$  è l'insieme degli  $N$  nodi,  $L$  è l'insieme di  $L$  coppie di elementi di  $N$  e  $W = \{w_1, w_2, \dots, w_K\}$ , con  $(i, j) = 1, 2, \dots, K$ , è l'insieme dei numeri reali associati a ciascun arco, detti *pesi*. Un grafo pesato è denotato con  $G^W$  [4, 5]. Se non indicato esplicitamente, una definizione di una grandezza caratteristica è equivalente per entrambi i tipi di grafo (non pesato o pesato), mentre in caso risulti necessaria, la distinzione verrà opportunamente evidenziata.

### 1.1.3 Proprietà dei nodi: matrice di adiacenza, grado di un nodo e assortatività

La rappresentazione grafica di una rete non è spesso la soluzione migliore per rappresentare un network complesso, specie se è composto da un grande numero di nodi. Inoltre, le reti cerebrali in particolare vanno estrapolate dai dati di imaging o dalla trattografia anatomica ossia non esistono di per se stesse come grafico di tipo visuale. Per questo, è spesso utilizzata una *rappresentazione matriciale dei grafi*, a partire

dalla quale è possibile ricavare uno tra i parametri fondamentali che caratterizzano i nodi, ossia il *grado* o *forza*, a seconda che rispettivamente si analizzi un grafo non pesato o uno pesato. Per un grafo binario  $G(N,K)$ , si definisce *matrice di adiacenza*  $A$  una matrice quadrata  $N \times N$  i cui elementi  $a_{ij}$  sono del tipo:

- $a_{ij} = 1$  se il link  $l_{ij}$  esiste;
- $a_{ij} = 0$  se il link  $l_{ij}$  non esiste;
- $a_{ii} = 0$  cioè la diagonale è costituita da soli zeri.

È evidente che per un grafo non orientato,  $A$  è una matrice simmetrica: essa rappresenta gli archi che connettono due nodi della rete come suoi elementi unitari.

Si definisce *grado*  $k_i$  del nodo  $i$  [1] il numero di archi incidenti nel nodo  $i$  ossia il numero di archi connessi a tale nodo.

$$k_i = \sum_{j \in N} a_{ij} \quad (1.1)$$

Se il grafo è orientato,  $k_i$  è detto *grado totale* del nodo  $i$  ed è dato dalla somma di due termini: uno dovuto agli archi uscenti dal nodo  $i$ , l'altro dovuto agli archi entranti in  $i$ .

$$k_i = k_i^{out} + k_i^{in} \quad (1.2)$$

dove  $k_i^{out} = \sum_{j \in N} a_{ij}$  e  $k_i^{in} = \sum_{j \in N} a_{ji}$ .

Per un grafo pesato  $G(N,K,W)$ , la matrice di adiacenza è sostituita dalla *matrice dei pesi*  $M$ , una matrice  $N \times N$  in cui l'elemento generico  $w_{ij}$  è il peso attribuito all'arco che collega il nodo  $i$  col nodo  $j$ , mentre  $w_{ij} = 0$  se  $i$  e  $j$  non sono connessi. Per convenzione,  $w_{ii} = 0$  per ogni  $i$ . Spesso si indica  $M$  con il nome di matrice di adiacenza in analogia al caso non pesato. Per lo studio di reti cerebrali, si assume che il grafo pesato non sia orientato e che i pesi siano tutti positivi, ossia che le interazioni non siano di natura distruttiva, ma sempre di natura costruttiva, dato che interazioni di tipo distruttivo non si osservano nelle reti di questo tipo mentre descrivono bene i fenomeni di altri tipo di reti quali quelle sociali. Di conseguenza,

la matrice  $M$  è simmetrica e  $w_{ij} = w_{ji} \geq 0$ . Per i grafi pesati, il concetto di grado è generalizzato dalla nozione di *forza* di un nodo  $i$  [4], definita come la somma di tutti i pesi degli archi connessi al nodo  $i$ .

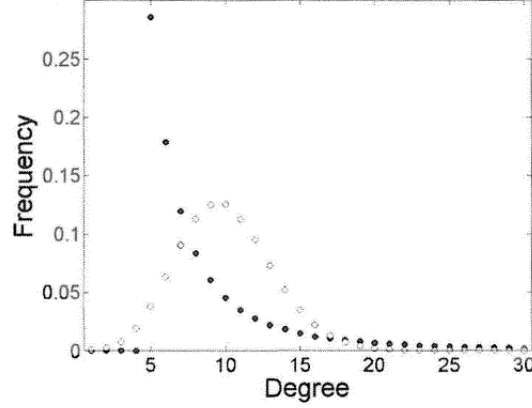
$$s_i = \sum_{j \in N} w_{ij} \quad (1.3)$$

Tale definizione amplia quella di grado poiché contiene sia l'informazione di quanti archi incidono in  $i$ , sia del loro peso.

Sia il grado che la forza dei nodi di una rete sono analizzati similmente (d'ora in avanti, si parlerà solo del grado data la somiglianza dell'approccio) attraverso la *distribuzione dei gradi*  $P(k)$ : tale funzione rappresenta la frazione dei nodi aventi un certo grado  $k$  o, in modo equivalente, la probabilità che un nodo selezionato a caso tra gli  $N$  nodi della rete abbia grado  $k$ . In caso di grafi orientati, si può considerare una distribuzione di probabilità per il grado  $k^{in}$  ed una per il grado  $k^{out}$ . La distribuzione  $P(k)$  può essere rappresentata tramite istogrammi dai quali estrapolare poi precisi andamenti funzionali, fermo restando che  $k$  può assumere solo valori interi positivi. Essa è la caratterizzazione topologica più basilare di un grafo e fornisce informazioni strutturali: infatti, dall'andamento di essa è possibile stimare se una rete possiede nodi con un grado approssimativamente uguale, piccati intorno ad un certo grado medio, oppure se essi hanno gradi distribuiti su un ampio range di valori, sintomo di una architettura radicalmente differente. Si definisce grado medio della rete il primo momento della distribuzione del grado  $P(k)$ :

$$\langle k^n \rangle = \sum_k k^n P(k) \quad (1.4)$$

Nel caso delle reti reali, tuttavia, la distribuzione  $P(k)$  non chiarisce completamente le proprietà statistiche della rete ed è necessario stimare la *correlazione* tra il grado di due nodi vicini, per spiegare il motivo dell'esistenza di connessioni preferenziali tra nodi di pari o diverso grado: infatti, un grande numero di networks è *correlato* cioè la probabilità che un nodo di grado  $k$  sia connesso ad un altro di grado  $k'$  dipende da  $k$ . Di conseguenza, è necessario introdurre la probabilità condizionata



**Figura 3:** Distribuzioni del grado  $P(k)$  di una rete ad invarianza di scala (puntini neri) e di una rete random (punti vuoti). Grafico medio per 10 reti, con 100000 nodi e grado medio pari a 10. Modificato da [3].

$P(k'|k)$  = probabilità che un nodo di grado  $k$  sia connesso ad uno di grado  $k'$ , che soddisfa le seguenti proprietà [6]:

1. condizione di normalizzazione:  $\sum_{k'} P(k'|k) = 1$
2. condizione di equilibrio “dettagliato”: all’equilibrio, ciascun processo elementare deve essere controbilanciato dal processo contrario, ossia devono essere equiprobabili. In termini quantitativi:  $kP(k'|k)P(k) = k'P(k|k')P(k')$ .

Per grafi non correlati,  $P(k'|k)$  non dipende da  $k$  e quindi partendo dalla condizione di equilibrio, tenendo conto della normalizzazione, sommando su  $k$  si ottiene:  $P(k'|k) = \frac{k'P(k')}{\langle k \rangle}$ . Per i grafi correlati invece, la stima di  $P(k'|k)$  è problematica: tale problema viene aggirato definendo il *grado medio dei nodi adiacenti al nodo  $i$* .

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in N_i} k_j = \frac{1}{k_i} \sum_{j=1}^N a_{ij} k_j \quad (1.5)$$

Nella Eq. (1.5), la somma è effettuata sugli elementi dell’insieme  $N_i$  dei nodi adiacenti al nodo  $i$  e costituisce una vera e propria media, essendo pesata sul grado

$k_i$  del nodo  $i$  che rappresenta anche il numero dei nodi vicini. In analogia a tale definizione, è possibile calcolare il *grado medio dei nodi adiacenti ad un nodo di grado  $k$*   $k_{nn}(k)$ , numericamente ottenuto tramite la definizione precedente, ma formalmente definito in modo implicito dalla distribuzione di probabilità condizionata  $P(k'|k)$ , di cui rappresenta il primo momento [7].

$$k_{nn}(k) = \sum_{k'} k' P(k'|k) \quad (1.6)$$

In assenza di correlazione, si ha  $P(k'|k) = \frac{k' P(k')}{\langle k \rangle}$  e sostituendo nella Eq. (1.6), si ottiene  $k_{nn}(k) = \frac{\langle k^2 \rangle}{\langle k \rangle}$  ossia  $k_{nn}(k)$  è una costante indipendente da  $k$ . In presenza di correlazione invece,  $k_{nn}(k)$  è una funzione di  $k$  ed a seconda del suo andamento è possibile stimare un importante parametro detto *assortatività* del grafo [8]:

- se  $k_{nn}(k)$  cresce con  $k$ ,  $G$  è detto assortativo o con *assortatività positiva*: in questo tipo di grafi, i nodi tendono a legarsi preferenzialmente a nodi di grado simile;
- se  $k_{nn}(k)$  decresce con  $k$ ,  $G$  è detto disassortativo o con *assortatività negativa*: nodi con alto grado tendono a connettersi con nodi di grado più basso.

#### 1.1.4 Proprietà dei collegamenti tra nodi: walk, path, distanza, connessione di un grafo

I nodi all'interno di un network possono essere connessi da singoli archi, come nel caso di nodi adiacenti, o da più archi.

Si definisce *percorso* (*walk*) dal nodo  $i$  al nodo  $j$  una sequenza alternata di nodi ed archi che comincia con  $i$  e finisce con  $j$  in cui gli archi possono essere percorsi più volte, consentendo che i nodi appartenenti al percorso possano essere visitati più di una volta.

Si definisce *cammino* (*path*) un percorso in cui ciascun arco è percorso solo una volta: di conseguenza, anche i nodi interni al cammino saranno percorsi solo una

volta. Un cammino può connettere un nodo con se stesso: un cammino chiuso costituito da almeno 3 archi è detto *ciclo* [1].

Per poter misurare e confrontare i cammini (ed i percorsi) tra due nodi, si può stabilire la nozione di *lunghezza*, tenendo ben presente che si tratta di una caratterizzazione topologica e non metrica, che non si riferisce alla separazione spaziale tra i nodi. In un grafo binario, la lunghezza  $L_{ij}$  di un cammino è definita come il numero di archi che lo costituiscono; di particolare interesse per le proprietà macroscopiche del grafo che ne derivano, è il *cammino di minima lunghezza* tra due nodi  $i$  e  $j$ , ossia quello costituito dal minor numero di links: la sua lunghezza è detta *distanza* tra  $i$  e  $j$ ,  $d_{ij}$ . È evidente che il cammino di minima lunghezza è il cammino più efficiente che si può stabilire tra due nodi, in quanto permette che l'informazione tra i due nodi estremi viaggi nel modo più rapido possibile.

In un grafo pesato, la lunghezza di un arco  $L_{ij}$  è definita come il reciproco del suo peso,  $L_{ij} = \frac{1}{w_{ij}}$ , e la somma dei reciproci dei pesi degli archi che costituiscono il cammino è la lunghezza del cammino [5]. In questa definizione, data la dipendenza dai pesi dei singoli archi, appare evidente che il cammino di minima lunghezza può non essere quello costituito dal numero di minimo di archi. La definizione si riduce al caso binario considerando ciascun arco di lunghezza unitaria  $L_{ij} = 1$  e di conseguenza  $d_{ij} = \min\{L_{ij}\}$ .

Le distanze tra i nodi possono essere organizzate sottoforma di una *matrice delle distanze*  $D$ , i cui elementi sono appunto le distanze  $d_{ij}$  tra i nodi. Il massimo assoluto di  $d_{ij}$  è detto *diametro* e si indica con  $Diam(G)$ . Un grafo  $G$  tale che per ogni coppia di due suoi nodi distinti  $i$  e  $j$  esiste almeno un cammino che li collega si dice *connesso* altrimenti è detto *non connesso*.

## 1.2 Misure caratteristiche delle reti

La maggior parte delle misure che descrivono le reti complesse posso essere derivate dall'analisi della matrici di adiacenza, dalla matrice delle distanze. A seconda delle proprietà che quantificano, le misure che caratterizzano una rete si posso distinguere in [3]:

1. *misure di segregazione locale*: stabiliscono l'influenza sulla connettività della rete da parte di insiemi di nodi adiacenti rispetto a un dato nodo e quantificano la tendenza alla formazione di regioni popolate da nodi fortemente interconnessi tra loro e poco connessi con gli altri nodi della rete;
2. *misure di integrazione globale*: quantificano l'efficienza della rete nel trasmettere informazioni a livello globale, anche tra nodi distanti, nell'accezione prima definita;
3. *misure di influenza di singoli nodi*: analizzano il ruolo preponderante che alcuni particolari nodi assumono all'interno della rete, classificandoli a seconda della centralità che essi hanno nella gestione della trasmissione del segnale tra una regione ed un'altra della rete.

### 1.2.1 Misure di segregazione locale: clustering, motivi e modularità

Questi parametri stimano quanto una rete è organizzata in insiemi di nodi adiacenti densamente accoppiati l'un l'altro: nodi altamente connessi infatti promuovono uno scambio più efficiente e rapido di informazione tra di loro, stabilendo di fatto connessioni più “forti” e formando dei *cluster* (“gruppi o ammassi”), ben localizzati e separati uno rispetto all'altro. Si può dire che il contributo del singolo nodo all'interno della rete è amplificato quando esso opera in stretta collaborazione con i suoi vicini: nell'ambito particolare delle reti cerebrali inoltre, questa modalità aggregativa dei nodi assume un preciso senso funzionale, poiché permette la formazione di aree coerenti nel cervello, ossia con una precisa funzionalità, come verrà mostrato nel capitolo 4. Stime per la tendenza a formare clusters sono il *coefficiente di clustering*, la *transitività*, la rilevazione di *motivi* e la *modularità*.

La misura più elementare di segregazione locale è il *coefficiente di clustering di un nodo  $i$*  [9]. Esso stima la probabilità che due nodi  $j$  e  $m$  adiacenti ad un nodo  $i$  di un grafo  $G$  siano adiacenti fra loro: in altri termini, è una stima del numero di

triangoli in cui il nodo  $i$  è uno dei vertici. È definito come

$$c_i = \frac{2e_i}{k_i(k_i - 1)} = \frac{\sum_{j,m} a_{ij}a_{jm}a_{mi}}{k_i(k_i - 1)} \text{ con } 0 \leq c_i \leq 1 \quad (1.7)$$

dove:

- i termini  $a_{ij}$  sono gli elementi della matrice di adiacenza, che possono essere anche nulli nel caso di nodi adiacenti ad  $i$  non connessi tra loro;
- $e_i$  è il numero effettivo degli archi all'interno del grafo  $G_i$ , detto anche sottografo indotto da  $N_i$ , insieme dei nodi adiacenti al nodo  $i$ .
- il numero  $\frac{k_i(k_i - 1)}{2}$  rappresenta il numero massimo di archi che possono essere presenti nel sottografo  $G_i$ , se  $i$  è un nodo di grado  $k_i$ .

Dalla definizione del coefficiente di clustering  $c_i$ , del singolo nodo  $i$ , è possibile ricavare un parametro macroscopico che caratterizza l'intera rete, ossia il *coefficiente di clustering del grafo*  $C$  definito come la media degli  $c_i$ , fatta su tutti gli  $N$  nodi del grafo  $G$ .

$$C = \frac{1}{N} \sum_{i \in N} c_i \text{ con } 0 \leq C \leq 1 \quad (1.8)$$

Anche  $C$  rappresenta quindi una probabilità ossia la frazione di triangoli esistente nel network, scelti a caso un nodo  $i$  e due suoi nodi adiacenti, rispetto a triplette di nodi formati dal nodo  $i$  e da due suoi vicini rispetto ai quali non si fa alcuna ipotesi di connessione.

Per grafi pesati, è possibile modificare la definizione di coefficiente di clustering del nodo tenendo conto dei pesi dei nodi adiacenti ad  $i$  e della forza del nodo  $i$ -esimo, ottenendo la naturale estensione della definizione precedente [4].

$$c_i^w = \frac{1}{s_i(k_i - 1)} \frac{\sum_{j,m} (w_{ij} + w_{im}) a_{ij} a_{jm} a_{mi}}{2} \text{ con } 0 \leq c_i^w \leq 1 \quad (1.9)$$

Il coefficiente  $c_i^w$  considera sia il numero di triangoli chiusi presenti nel nell'insieme dei nodi adiacenti ad  $i$ , nel momento in cui tutti e 3 i coefficienti  $a_{ij}$ . sono non nulli

e quindi uguali ad 1, sia il loro peso totale  $w_{ij} + w_{im}$  rispetto alla forza  $s_i$ , del nodo  $i$ . Il termine a denominatore  $s_i(k_i - 1)$  è un fattore di normalizzazione in modo che  $c_i^w$  abbia ancora un significato probabilistico. Si nota come sostituendo ad  $s_i$  il suo corrispettivo per grafi binari, ossia il grado  $k_i$  e considerando i pesi  $w_{ij}$  unitari, si ritrova la definizione di  $c_i$  data precedentemente per i grafi non pesati.

Nella definizione di  $c_i$ , e di conseguenza in quella di  $C$  (Eq. 1.8), si nota la dipendenza dal grado del nodo  $k_i$ : nodi di basso grado possono in qualche modo falsare la misura del coefficiente di clustering. Per questo, la misura di  $C$  è indicata in grafi con una distribuzione del grado omogenea, piccata attorno ad un grado medio. Per ovviare al problema nella generalità dei casi, una misura alternativa della proprietà di clustering di una rete complessa è fornita dalla *transitività*  $T$  [10]: analogamente a  $C$ , questo coefficiente stima la presenza dei triangoli, prescindendo dal grado di uno dei 3 nodi che lo formano, come la frazione di triangoli rispetto alla totalità dei triplette formati da 3 nodi connessi fra loro. Esso è definito come:

$$T = \frac{3 * \sum_{i \in N} t_i}{\sum_{i \in N} tr_i} \quad (1.10)$$

dove:

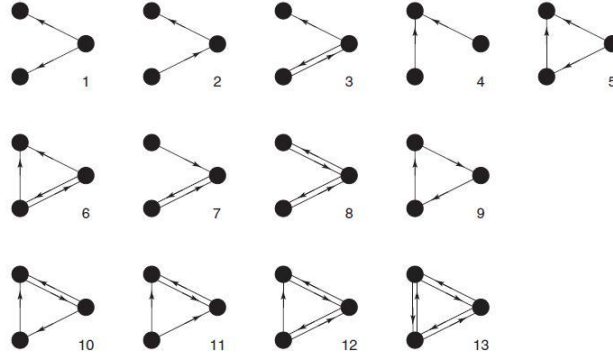
- il numeratore rappresenta, a parte il fattore 3, il numero dei triangoli in  $G$ , dato dalla somma del numero di triangoli attorno a ciascun nodo  $i$ :

$$t_i = \frac{1}{2} \sum_{j, m \in N} a_{ij} a_{jm} a_{mi}$$

- il denominatore rappresenta il numero dei triplette connessi in  $G$ , dato dalla somma del numero dei triplette che coinvolgono ciascun nodo  $i$ :

$$tr_i = \frac{1}{2} \sum_{j, m \in N} a_{ij} a_{im}$$

Il fattore 3 nella formula è dovuto al fatto che ciascun triangolo formato da 3 vertici fornisce 3 triplette di nodi connessi, ciascuno centrato su un vertice. Da un certo punto di vista, può essere vista come stima della validità di una “proprietà



**Figura 4:** I 13 motivi possibili composti da 3 nodi connessi da link orientati [1].

transitiva” rispetto alla relazione che lega due nodi: nelle reti sociali, ad esempio, è la probabilità che due individui che hanno un amico in comune siano mutualmente amici fra loro.

All’interno dei singoli cluster locali di nodi, i tipi di connessione che essi possono stabilire possono essere diversi: lo studio di numerosi casi di reti reali mostra che l’organizzazione dei nodi all’interno dei cluster non è casuale ma segue dei *motivi* ben definiti [11], sottografi che formano degli schemi topologici di interconnessione ricorrenti in una data rete.

Un motivo  $M$ , sottografo di  $G$ , può essere visto come una sorta di “elemento circuitale” di base in cui la rete è scomponibile: la “ricorrenza” di un motivo  $M$  è indice innanzitutto di una struttura non randomica e specializzando il discorso circa il tipo di rete, la prevalenza di un motivo piuttosto che un altro potrebbe essere indice di una specializzazione funzionale, come sarà mostrato dopo nell’analisi delle reti cerebrali. È necessario specificare il senso del termine “ricorrenza” relativamente ad un motivo: per stabilire se un motivo  $M$  è caratteristico del grafo  $G$  oggetto di studio, ossia è presente in  $G$  in modo “significativamente alto”, è necessario confrontare il grafo  $G$  con un insieme di altri grafi  $G'$ , che rappresentano la versione randomizzata di  $G$ . Se  $G$  è un grafo con  $N$  nodi,  $K$  link caratterizzato da una certa distribuzione del grado  $P(k)$ , allora  $G'$  è un grafo con stesso numero di nodi  $N$ , stesso numero di link  $K$  e stessa distribuzione  $P(k)$ , ma in cui i link sono assegnati in maniera casuale ai nodi. Il grafo  $G'$  serve come “ipotesi nulla”, ossia come grafo-test in cui l’emergenza

di un motivo piuttosto che di un altro è dovuta solo a fluttuazioni statistiche e non ad una effettiva ricorrenza strutturale. La significatività statistica di un motivo  $M$  è data dal valore- $Z$ , definito come [12]:

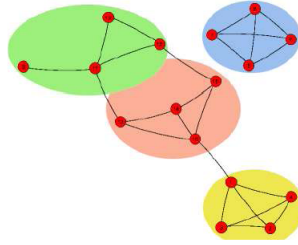
$$Z_M = \frac{n_M - \langle n_M^{rand} \rangle}{\sigma_M^{rand}} \quad (1.11)$$

dove

- $n_M$  è il numero di volte in cui il sottografo o motivo  $M$  appare in  $G$
- $\langle n_M^{rand} \rangle$  è il numero medio di volte in cui  $M$  compare nei grafi random  $G'$  mentre  $\sigma_M^{rand}$  è la deviazione standard dalla media.

Reti reali con alto coefficiente di clustering mostrano spesso una organizzazione in comunità locali di cluster di nodi o *moduli* di nodi fortemente coesi al loro interno e segregati l'uno rispetto all'altro: questa osservazione empirica suggerisce l'introduzione di una misura di *modularità* di una rete [13], che stimi in che misura la rete è composta da moduli in cui al loro interno numerosi archi connettono i nodi, mentre pochi archi connettono nodi tra moduli diversi.

La definizione di una simile misura non è banale poiché è legata al modo in cui è ricercata l'esistenza di moduli: questo tipo di risultato è ottenuto tramite l'applicazione di algoritmi basati sull'idea del partizionamento della rete in step successivi, valutando di volta in volta se la partizione in moduli ottenuta è coerente con quella realmente osservata: questa valutazione di compatibilità è effettuata tramite funzioni test che raggiungono valori di massimo in corrispondenza della partizione migliore. È necessario puntualizzare che è proprio l'ottimizzazione con la rete reale oggetto di studio a garantire che la modularità trovata non sia legata al metodo di misura, ma è effettivamente una proprietà intrinseca della rete. Esistono diversi algoritmi di partizionamento, che differiscono oltre che la filosofia anche per il loro range di applicazione [1]: ciò costituisce una limitazione per quegli algoritmi con campo di azione ristretto come il *partizionamento spettrale* che soffre del grave handicap che è necessario conoscere a priori almeno una stima del numero di comunità presenti



**Figura 5:** Moduli di un grafo

nella rete e spesso ciò non è noto, a maggior ragione nel caso di reti non direttamente ricavabili in modo visivo come quelle cerebrali. Un'altra serie di algoritmi, di portata più generale in quanto non richiedono la conoscenza a priori del numero indicativo di moduli è quella degli *algoritmi di clustering gerarchico*, in cui si considerano una serie di partizioni possibili che vanno dall'esistenza di un singolo cluster contenente tutti gli  $N$  nodi ad  $N$  clusters costituiti ciascuno da ogni singolo nodo. Si distinguono così algoritmi di clustering gerarchico di tipo agglomerativo, in cui si fondono gli  $N$  nodi in gruppi più grandi, o viceversa di tipo divisivo, in cui si separa il cluster di  $N$  oggetti in gruppi sempre più ristretti. A ciascuno step, è valutata la distanza topologica fra i cluster ottenuti: il limite di questi algoritmi è che non è definito con precisione quale sia lo step per cui si ottiene la partizione migliore, essendo la decisione di tipo euristico.

L'algoritmo più efficace e di portata più generale per analizzare la modularità di una rete è quello proposto da Girvan e Newman [14], per questo chiamato anche *metodo GN*: è un metodo divisivo basato sulla ricerca e progressiva rimozione degli archi con il più grande valore di betweenness, un parametro di integrazione globale che verrà discusso nel prossimo paragrafo, che indica la centralità di un arco, cioè la sua tendenza nel prendere parte ai cammini attraverso cui passa l'informazione: in poche parole, la sua influenza. Dal momento che gli archi che si estendono tra due nodi appartenenti a due moduli differenti sono quelli più influenti per la coesione della rete, è verosimile credere che siano anche quelli con il più alto valore di betweenness: rimuovendoli ricorsivamente, si ottiene la separazione della rete in comunità. In generale, l'algoritmo segue tale schema:

1. calcolo del valore di betweenness per ciascun arco;
2. rimozione degli archi con il valore più alto;
3. analisi dei componenti della rete così ottenuti;
4. ripetizione dei passi precedenti fino a che ogni arco è rimosso ed il sistema è frammentato in N nodi non connessi.

L'algoritmo fornisce una suddivisione in moduli ad ogni iterazione, generando quindi una suddivisione gerarchica della rete: per stabilire qual è la suddivisione migliore per una data rete ossia dove troncare l'albero gerarchico delle suddivisioni, si usa il parametro  $Q$ , detto *modularità* che può quindi essere definito in modo preciso, avendo fissato l'algoritmo di ricerca. Qualitativamente, una buona suddivisione in un numero  $n$  di moduli è quella in cui la maggior parte degli archi cadono all'interno degli  $n$  moduli, mentre solo pochi links connettono nodi appartenenti a comunità diverse; quantitativamente, ciò si può formalizzare introducendo una matrice  $n \times n$  i cui elementi  $e_{ij}$  sono la frazione di tutti gli archi nella rete che connettono nodi nel modulo  $i$  a nodi nel modulo  $j$ . La traccia della matrice  $E$ ,  $TrE = \sum_i e_{ii}$ , è la frazione di tutti gli archi che collegano nodi all'interno dello stesso modulo, per ogni modulo  $i$ -esimo, mentre la somma degli elementi di una sua riga  $i$ -esima fissata (analogamente di una colonna),  $a_i = \sum_j e_{ij}$ , rappresenta la frazione degli archi che si connettono a nodi tutti appartenenti al modulo  $i$ . La modularità  $Q$  è quindi definita come:

$$Q = \sum_i (e_{ii} - a_i^2) = TrE - \|E^2\| \quad (1.12)$$

dove  $E^2$  rappresenta la somma degli elementi della matrice  $E^2$ . Di conseguenza,  $Q$  misura il grado di correlazione tra la probabilità che esista un arco che connette due nodi ( $E^2$ ) e il fatto che questi due nodi appartengano allo stesso modulo ( $TrE$ ).  $Q$  varia da un minimo di 0, in caso di un grafo privo di moduli (es. un grafo random), e tende ad un massimo di 1, in caso di una forte struttura modulare, indicando una grande diversità tra il numero di link interni ad una comunità e quelli che legano comunità differenti. Il valore di  $Q$  è computato a ciascuna iterazione, cioè per diversi

valori del numero di suddivisioni  $n$ : picchi dei valori  $Q$  verso 1 durante l'applicazione dell'algoritmo indicano suddivisioni ottimali. Il metodo GN è largamente impiegato in numerosi studi di reti, comprese quelle cerebrali: l'aspetto fondamentale dell'algoritmo, ossia il calcolo della betweenness di ogni arco ad ogni iterazione costituisce però anche il suo limite. Infatti, è necessaria una potenza di calcolo molto elevata, che richiede un tempo dell'ordine del prodotto  $K * N$ . Gli hardware attualmente disponibili permettono l'applicazione del metodo per reti o porzioni di reti composte da nodi intorno a poche migliaia.

Le misure di segregazione possono spesso sovrapporsi nel senso che forniscono indicazioni simili, legate tra loro. Ad esempio, un altro coefficiente di clustering per una rete è spesso accompagnato da un grande valore della significatività  $Z$  per una determinata classe di motivi ma ciò non è sorprendente alla luce del fatto che si può interpretare il coefficiente  $C$  come la frazione di triangoli rispetto ai triplette di nodi connessi, ossia di motivi costituiti da 3 nodi completamente connessi. Viceversa però, può anche accadere che una misura di segregazione non implichi un'altra: un grande valore di modularità spesso implica un alto coefficiente di clustering, ma non è vero il viceversa. Basti pensare ai grafi regolari, in cui  $C$  è molto alto ma i moduli sono completamente assenti. In conclusione, la ridondanza tra le diverse misure di segregazione locale è solo parziale e a ciascuna compete una informazione specifica circa le proprietà locali della rete.

### **1.2.2 Misure di integrazione globale: cammino caratteristico ed efficienza globale**

Le misure di integrazione locale catturano la capacità della rete di stabilire pattern di connessioni che superino i confini dei moduli, o ad un livello più fine, dei cluster di nodi stessi, per favorire l'efficienza della rete nello scambio dei segnali e delle informazioni tra le sue varie componenti. Si può immaginare come una rete possa soddisfare a due richieste strutturali diametralmente opposte come la segregazione e l'integrazione, operando un bilanciamento tra queste caratteristiche, volto alla massimizzazione della funzionalità della rete come una entità coesa. Le misure carat-

teristiche dell'integrazione globale di una rete sono basate sulle nozioni di lunghezza di un cammino e di distanza.

La *lunghezza caratteristica di un cammino*  $L$  di un grafo  $G$ , detta anche più semplicemente *cammino caratteristico* [9] è la media delle distanze  $d_{ij}$  tra ogni coppia di nodi  $i$  e  $j$  scelti tra gli  $N$  nodi che costituiscono il grafo  $G$ .

$$L = \frac{1}{N(N-1)} \sum_{i,j \in N, i \neq j} d_{ij} \quad (1.13)$$

$L$  è una misura della separazione (topologica e non metrica) tra due nodi del grafo scelti a caso: un piccolo valore di  $L$  implica un breve cammino caratteristico ossia in media ciascun nodo può essere raggiunto a partire ogni altro nodo attraverso cammini brevi, costituiti da pochi archi nel caso di grafi binari. Il valore di  $L$  assume una maggiore significatività se confrontato con il cammino caratteristico  $L'$  dell'equivalente grafo random  $G'$ , che anche in questo caso ha il ruolo di valore di riferimento: tanto più  $L$  si discosta da  $L'$ , tanto più emerge una architettura non randomica del grafo, che ha i suoi effetti nell'ottimizzazione del cammino caratteristico. Dalla definizione di  $L$ , tuttavia, emerge un problema di natura computazionale: se due nodi  $i$  e  $j$  del grafo non sono connessi, allora  $d_{ij} \rightarrow \infty$  e quindi  $L$  diverge. Per ovviare a ciò, è possibile in prima istanza, limitare il calcolo di  $L$  al modulo connesso più grande. L'approccio che si preferisce è invece definire un'altra misura, fortemente correlata al cammino caratteristico, detta *efficienza globale*  $E$  del grafo  $G$  [15], media armonica delle distanze  $d_{ij}$ .

$$E = \frac{1}{N(N-1)} \sum_{i,j \in N, i \neq j} \frac{1}{d_{ij}} \quad (1.14)$$

È evidente che per due nodi  $i$  e  $j$  non connessi la quantità  $1/d_{ij}$  è nulla, evitando il problema della divergenza. La quantità  $E$  è un indice della capacità della rete di sopportare e gestire un flusso ingente di informazioni e segnali: tanto più grande è  $E$ , tanto minore è il cammino caratteristico, tanto più è efficiente la rete nello "smaltire" il traffico di dati nel senso che consente all'informazione di viaggiare da un nodo qualunque ad un altro nel modo più rapido possibile, secondo i cammini

più brevi. Una rete con alta efficienza quindi dispone i suoi nodi a breve distanza topologica l'uno dall'altro, consentendo loro di interagire più direttamente per una migliore connettività globale.

### 1.2.3 Misure di centralità dei nodi: closeness e betweenness, hubs

Nelle reti reali, emerge l'esistenza di nodi più influenti di altri nella funzionalità della rete: se questo è un concetto non nuovo per gli archi, per i quali è messo in evidenza e quantificato in termini di peso di un arco, è invece una nuova prospettiva grazie alla quale è possibile classificare i nodi che mostrano di avere un maggiore rilievo per la robustezza e l'efficienza della rete. I nodi di “maggiore importanza” derivano questo attributo dal fatto che essi sono spesso più altamente connessi rispetto agli altri perciò, dovendo sostenere un traffico maggiore di informazione, la loro rimozione o danneggiamento risultano catastrofici. I nodi con simili caratteristiche sono detti *hub* [3], sebbene il termine di per sé non sia ben definito. I nodi candidati ad essere possibili hubs sono classificati in base a diverse caratteristiche, quali il grado, la partecipazione nei cammini che collegano moduli differenti e le misure di *centralità*.

L'indicatore di più facile interpretazione è il grado: questo parametro può essere particolarmente significativo in quei network che hanno una distribuzione del grado  $P(k)$  molto estesa, spalmata su un ampio range e non piccata attorno ad un grado medio. In queste reti infatti l'esistenza di pochi nodi con alto grado è il collante essenziale per il mantenimento della compattezza della rete; diversamente, in reti con una distribuzione  $P(k)$  omogenea, il grado non è un indicatore particolarmente rilevante per la determinazione di un hub. Un altro stimatore della presenza di un hub può essere considerato il *coefficiente di partecipazione di un nodo* [16], che quantifica nell'ambito di reti modulari il numero di connessioni tra nodo di alto grado all'interno di un modulo  $i$  e un altro nodo di alto grado appartenente ad un altro modulo. Esso è definito come:

$$P_i = 1 - \sum_{s=1}^{N_M} \frac{k_{is}}{k_i} \quad (1.15)$$

dove:

- $N_M$  indica il numero totale di moduli identificati nel network;
- $k_i$  è il grado del nodo  $i$ ;
- $k_{is}$  è il numero di archi che connettono il nodo  $i$  a ciascuno dei moduli  $s$ .

Un nodo di alto grado con un alto coefficiente di partecipazione,  $P \geq 0.3$ , promuove l'interazione intermodulare e quindi l'integrazione della rete ed è quindi un candidato *hub connettore*, mentre un nodo di alto grado con un basso coefficiente di partecipazione,  $P < 0.3$  partecipa principalmente alle connessioni all'interno del modulo a cui appartiene, favorendo la coesione della comunità di cui è parte, ed è quindi un possibile *hub provinciale*.

Altre misure quantitativamente più rilevanti per l'individuazione di possibili hubs sono le misure di centralità [17]: possono essere di due tipi a seconda che ci si riferisca alle distanze topologiche fra i nodi o alla capacità di un nodo di “prendere parte” al maggior numero di cammini possibili. La “vicinanza” o *closeness* di un nodo  $i$  stima la centralità di un nodo in base a quanto esso è prossimo a tutti gli altri nodi della rete. Considerata una rete con  $N$  nodi, tale misura è definita come l'inverso della distanza media di un nodo  $i$  rispetto a tutti gli altri nodi.

$$cl_i = \frac{1}{\frac{1}{N-1} \sum_{j \in N, j \neq i} d_{ij}} = \frac{N-1}{\sum_{j \in N, j \neq i} d_{ij}} \quad (1.16)$$

Se la distanza media è bassa, il nodo  $i$  è “vicino” ossia è centrale rispetto a tutti gli altri e quindi è in una posizione di “potere” sia perché può esercitare una influenza più diretta sugli altri nodi sia perché gestisce una quantità di informazioni maggiore, rendendo la rete sia particolarmente vulnerabile alla sua rimozione. La “capacità di intermediazione” o *betweenness* di un nodo invece stima la rilevanza di un nodo in base al numero di cammini di minima lunghezza che lo attraversano. Considerato un grafo  $G$  con  $N$  nodi, la *betweenness* di un nodo  $i$  è definita come la frazione di cammini di minima lunghezza che connettono due nodi arbitrari  $j$  e  $k$  che passano

per  $i$  rispetto al numero totale di cammini di minima lunghezza che connettono  $j$  e  $k$ .

$$b_i = \sum_{j,k \in N} \frac{n_{jk}(i)}{n_{jk}} \quad (1.17)$$

dove

- $n_{jk}(i)$  è il numero di cammini di minima lunghezza tra  $j$  e  $k$  passanti per  $i$ ;
- $n_{jk}$  è il numero totale di cammini di minima lunghezza tra  $j$  e  $k$ .

La definizione di betweenness data per i nodi è applicabile anche agli archi, dove la lettera  $i$  indicherà un ben fissato arco: questa misura è quella computata ad ogni iterazione del metodo GN per individuare la suddivisione in moduli ottimale di una rete [14]. Un nodo o un arco con alta betweenness è più importante rispetto ad altri poiché si trova all'intersezione di numeri cammini brevi e quindi si trova a gestire un alto tasso di informazione.

## 2 Topologia delle reti e modelli

### 2.1 Modelli ed architetture di reti ideali: reti random e reti regolari

Nell'ambito della teoria dei grafi, i primi studi e misurazioni delle proprietà topologiche esposte nel precedente capitolo sono state eseguiti su grafi-modello ideali, di derivazione puramente matematica. Di conseguenza, già la loro genesi, svincolata dalla realtà, deve indurre a pensare che non saranno queste le architetture più pertinenti per descrivere le reti reali e che saranno necessari modelli meglio adattati: tuttavia, il loro studio è fondamentale sia perché è formalizzato in maniera elegante e compatta, sia perché essi forniscono possibilità di classificazione ed un metro di paragone per i network reali, fungendo da “ipotesi nulla” rispetto ad essi, relativamente ad una data proprietà. I due modelli ideali di fondamentale importanza sono i grafi *random* ed i grafi *regolari*.

#### 2.1.1 Grafi random o grafi ER

I grafi random sono stati i primi ad essere studiati e ad essere pensati come modelli dei grafi reali: sono detti anche grafi ER dal nome dei due matematici ungheresi Erdős e Rényi che nel 1959 per primi ne proposero un'elegante descrizione, partendo dall'assunzione che la connessione tra due nodi fosse casuale, ossia esistesse con una certa probabilità e quindi che non vi fosse un ordine nella distribuzione dei link tra i nodi. Di conseguenza, ignorando di fatto tutte le possibili leggi di natura che regolano la formazione dei sistemi complessi, i due matematici considerarono la realtà, ed i grafi che avrebbero dovuto rappresentarla, del tutto basati sul caso: questa si è mostrata come la via più semplice e compatta per fornire un'iniziale chiave interpretativa dell'interazione fra i vari componenti di un sistema complesso.

Un grafo random costituito da  $N$  nodi e  $K$  archi può essere generato secondo due modelli di sviluppo che coincidono nel limite di grandi  $N$ , a fissato grado medio  $\langle k \rangle$  [18]. Il primo modello dà origine ad un grafo random  $G_{N,K}^{ER}$  partendo da una situazione iniziale dove è presente solo un insieme di  $N$  nodi non connessi: coppie di

nodi scelti a caso sono collegate mediante un unico arco fino a che il numero di archi non sia pari a  $K$ . Il grafo così ottenuto è solo un singolo risultato tra tutte le possibili combinazioni di connessione tra coppie di nodi: di conseguenza il simbolo  $G_{N,K}^{ER}$  denota l'intero insieme statistico a cui esso appartiene che è necessario conoscere per la completa descrizione dei grafi ottenuti. In termini di rappresentazione matriciale, la completa conoscenza dell'insieme  $G_{N,K}^{ER}$  corrisponde alla completa conoscenza dell'insieme di tutte le matrici di adiacenza  $N \times N$  possibili con  $K$  elementi [19].

Il secondo modello invece genera un grafo random con  $N$  nodi e  $K$  link connettendo ciascuna coppia di nodi con una probabilità  $0 < p < 1$  uniforme per ogni coppia di nodi: l'ensemble di grafi creato da questa procedura, denotato  $G_{N,p}^{ER}$  differisce dal precedente  $G_{N,K}^{ER}$  in quanto contiene grafi con un numero di link variabile, non necessariamente pari a  $K$ : i grafi random con  $K$  archi esistono nell'insieme  $G_{N,p}^{ER}$  con una probabilità

$$p^K (1-p)^{\frac{N(N-1)}{2}-K} \quad (2.1)$$

dove:

1.  $p^K$  è la probabilità che esistano  $K$  link, data dal prodotto di  $p$  per se stesso  $K$ -volte: questo indica che il fatto che due link siano connessi è un evento indipendente e non influenza la probabilità che altri due nodi siano connessi; non a caso,  $p$  è uniforme.
2.  $(1-p)^{(\frac{N(N-1)}{2}-K)}$  indica la probabilità che le possibili altre coppie di nodi rimanenti, il cui numero è dato dalla differenza tra il loro numero massimo  $N(N-1)/2$  e il numero di archi desiderato  $K$ , non sia connesso. Infatti, se  $p$  è la probabilità che due nodi siano connessi,  $(1-p)$ , suo complemento ad 1, indica la probabilità che tale evento non si verifichi [19, 20].

Lo studio delle proprietà dei grafi random è stato effettuato in modo particolare sui grafi generati tramite il secondo modello, poiché i calcoli si rivelano più semplici, benché il primo modello sembra a primo acchito più pertinente del secondo per le applicazioni alle reti reali: tale affermazione è giustificata se si pensa allo finalità originaria per cui furono implementati i grafi random, cioè lo studio delle reti sociali.

Le proprietà strutturali dei grafi random variano in funzione della probabilità di connessione  $p$ , che raggiunge un valore critico  $p_c = 1/N$  in corrispondenza del quale il grafo ha grado medio  $\langle k_c \rangle = 1$  ossia è *mediamente* connesso. L'evoluzione strutturale del grafo al variare della probabilità  $p$  è riscontrabile soprattutto dalla nascita e dimensione dei moduli che compongono la rete:

1. se  $p < p_c$ , allora quasi sicuramente (ossia con probabilità tendente ad 1 con  $N$  tendente ad infinito), il grafo ha moduli di dimensioni non più grandi di  $O(\ln N)$ ;
2. se  $p = p_c$ , quasi sicuramente il più grande modulo connesso ha dimensione dell'ordine di  $O(N^{2/3})$ ;
3. se  $p > p_c$ , quasi sicuramente il grafo presenta un modulo di dimensioni  $O(N)$ : quasi sicuramente, sarà quindi un grafo connesso [18, 20].

La distribuzione del grado  $P(k)$  può essere ottenuta dalle seguenti considerazioni [21]. La probabilità che un nodo  $i$  abbia grado  $k_i$  è data dalla distribuzione binomiale:

$$P(k = k_i) = C_{N-1}^k p^k (1-p)^{N-1-k} \quad (2.2)$$

dove:

- $p^k$  è la probabilità che esistano  $k$  archi incidenti nel nodo  $i$ , ossia che esso sia connesso a  $k$  nodi;
- $(1-p)^{(N-1-k)}$  è la probabilità che non esistono i rimanenti  $N-1-k$  archi incidenti in  $i$ ;
- $C_{N-1}^k = \binom{N-1}{k}$  è il coefficiente binomiale che indica il numero di possibili modi in cui  $i$   $k$  nodi adiacenti a  $i$  possono essere scelti tra gli  $N-1$  disponibili, una volta fissato il nodo  $i$ .

Dato che in una network random i nodi sono tutti equivalenti dal punto di vista della probabilità di essere connessi, ciascuno di loro ha la stessa probabilità di avere un

dato grado  $k$  e quindi la probabilità che un nodo scelto a caso abbia grado  $k$  è data dalla distribuzione binomiale appena esaminata, valida per ogni  $i$ .

Per grafi di grandi dimensioni, ossia per grandi  $N$  e fissato grado medio  $\langle k \rangle = Np$ , la distribuzione binomiale è ben approssimata dalla Poissoniana, ragion per cui i grafi ER sono spesso chiamati grafi random poissoniani:

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \quad (2.3)$$

Appare evidente quindi che in questo tipo i nodi hanno un grado quasi uniforme: si dice che la distribuzione  $P(k)$  ha una scala caratteristica, definita dal grado medio sulla quale è evidentemente piccata, sia nel caso binomiale che poissoniano (Fig. 3). Inoltre, dato che i nodi sono connessi a coppie indipendentemente dal loro grado, ma tutti con la stessa probabilità  $p$ , i grafi random sono per definizione non-correlati: di conseguenza il grado medio dei nodi adiacenti ad un nodo di grado  $k$  scelto a caso,  $k_{nn}(k)$  e la probabilità condizionata  $P(k'|k)$  che un nodo di grado  $k$  sia connesso ad uno di grado  $k'$  sono entrambe quantità indipendenti da  $k$ .

Importanti informazioni giungono dal calcolo della principale misura di integrazione globale, il cammino caratteristico  $L$  e della principale misura di segregazione locale, il coefficiente di clustering  $C$ . Sia il diametro che il cammino caratteristico  $L$  mostrano la stessa dipendenza funzionale da  $N$ :  $L \frac{\ln N}{\ln \langle k \rangle} = \frac{\ln N}{\ln(pN)} \rightarrow 1$  se  $N \rightarrow \infty$  [9, 20]. Il coefficiente di clustering  $C$  è invece dato dalla probabilità  $p$  [9]: infatti, dato un nodo  $i$  di grado  $k$ , dal momento che  $p$  è la probabilità che due nodi qualsiasi siano connessi, allora ci saranno  $pk(k-1)/2$  archi a connettere i  $k$  nodi adiacenti ad  $i$ , rispetto ad un massimo di  $k(k-1)/2$  archi possibili, quando i  $k$  nodi adiacenti ad  $i$  sono tutti connessi tra di loro. Di conseguenza:

$$C = \frac{\frac{pk(k-1)}{2}}{\frac{k(k-1)}{2}} = p = \frac{\langle k \rangle}{N} \rightarrow 0 \text{ se } N \rightarrow \infty \quad (2.4)$$

In conclusione, i grafi random sono caratterizzati da un basso cammino caratteristico  $L$ , tale che nei networks di grosse dimensioni coppie di nodi scelti a caso sono connessi da cammini piuttosto brevi ( $L \rightarrow 1$ ) ma da un altrettanto basso livello di clustering

( $C \rightarrow 0$ ): è molto improbabile che due nodi adiacenti rispetto ad un nodo fissato siano anche adiacenti fra di loro.

### 2.1.2 Grafi regolari

I grafi regolari [22] sono un altro semplice modello ideale di grafo e a differenza dei grafi random sono caratterizzati da schemi ordinati di connessione fra i nodi: esempi possono essere gli anelli, le griglie o reticoli. La proprietà peculiare di tali grafi è che ogni nodo ha lo stesso grado  $k$  ossia ciascun nodo ha lo sempre lo stesso numero di nodi adiacenti. Per dedurne le proprietà emergenti conviene far riferimento agli anelli: un anello composto da  $N$  nodi è un insieme di nodi disposti circolarmente, in cui ciascun nodo è connesso simmetricamente ai suoi  $2m$  nodi vicini,  $m$  nodi in senso orario ed altrettanti  $m$  in senso antiorario, formando un totale di  $K = mN$  archi. La scelta del valore di  $m$  è arbitraria ed ovviamente deve essere  $m \ll N/2$ .

Sin dal grafico, appare evidente che le connessioni tra i nodi tendono ad essere “localmente dense”, cioè due nodi adiacenti tra loro tendono ad avere gli stessi vicini ossia alto livello di clustering, ma scelti due nodi a caso, la distanza fra loro può non essere piccola e il cammino di minima lunghezza può coinvolgere numerosi archi. Di conseguenza, un grafo regolare mostra proprietà diametralmente opposte ad uno random: alto coefficiente di clustering e grande cammino caratteristico.

## 2.2 Proprietà emergenti delle reti reali: proprietà di piccolo mondo ed invarianza di scala

Sia i grafi random che quelli regolari, sebbene siano stati ampiamente studiati in letteratura e consentano descrizioni analitiche complete ed eleganti non sono adatti a descrivere la maggior parte delle reti reali. Infatti, negli ultimi due decenni, lo studio della topologia dei sistemi complessi di origine più disparata, sfruttando il potente mezzo delle reti complesse, ha rivelato un aspetto essenziale: nonostante la varietà dei sistemi studiati, siano essi sociali, biologici, informatici, reti di comunicazione e trasporti, ciascuno di loro è caratterizzato dalle stesse proprietà topologiche:

1. cammino caratteristico relativamente piccolo;
2. alto coefficiente di clustering;
3. distribuzione del grado  $P(k)$  con lunghe code, cioè con valori  $P(k) \neq 0$  anche a grandi  $k$ ;
4. esistenza di motivi e struttura modulare.

La presenza contemporanea di queste proprietà non è riscontrata in nessuno dei due modelli di grafo ideale considerati: di conseguenza nasce l'esigenza di costruire modelli appropriati a descrivere le reti reali, che tengano conto di questi segni particolari. Le caratteristiche topologiche sopra citate seguono da proprietà più generali riscontrate nelle reti reali: l'effetto di “*piccolo mondo*” e le distribuzioni del grado  $P(k)$  ad *invarianza di scala*. Una menzione a parte è necessaria per la proprietà di modularità, di particolare importanza per le reti cerebrali.

### 2.2.1 Effetto di “piccolo mondo”

In molte reti reali, nonostante le loro grandi dimensioni (ossia il grande numero  $N$  di nodi) esiste un cammino relativamente breve tra ciascuna coppia di nodi scelti a caso: ciò è dovuto all'esistenza di archi che fungono da “*shortcuts*”, vere e proprie scorciatoie che collegano nodi appartenenti ad aree differenti di un network, rendendo così più rapida la diffusione dell'informazione o del segnale. Questa proprietà, detta quindi di “*piccolo mondo*”, è stata osservata inizialmente nelle reti sociali (l'esperimento di Milgram mette in luce tale caratteristica delle reti sociali, mostrando che due individui sono in media collegati da non più di 6 gradi di separazione) ma è stata riscontrata in molte altre tipologie di reti incluse quelle tecnologiche e biologiche: le stesse reti cerebrali estratte per diverse specie mostrano una struttura “small-world”, sebbene con la richiesta aggiuntiva della modularità.

Da un punto di vista matematico, l'effetto small-world è quantificato da un cammino caratteristico piccolo (cioè simile a quello ottenuto per un insieme di reti random con lo stesso numero di nodi e archi) che va al più logaritmicamente con  $N$ ,

$L \approx \ln N$ , associata alla presenza di un alto coefficiente di clustering [9]. Si nota che queste due caratteristiche appartengono la prima ai grafi random, la seconda a quelli regolari: da qui, un primo segnale di inefficienza dei modelli ideali per descrivere le reti reali. L'esistenza contemporanea di questi due segni fornisce quindi la definizione di rete di piccolo mondo, secondo Watts e Strogatz, che per primi ne hanno proposto un modello. Una quantificazione della proprietà di piccolo mondo è stata fornita da Humphries et al. [23], tramite la definizione di un “*indice di piccolo mondo*”: esso esprime il rapporto tra il coefficiente di clustering  $C$  e il cammino caratteristico di un grafo  $L$ , normalizzati ai corrispondenti valori di queste grandezze valutati per un grafo random ( $C_0$  ed  $L_0$ ) con lo stesso numero di nodi, archi e grado dei nodi. Il grafo random qui è utilizzato come “ipotesi nulla” relativamente alle due misure considerate.

$$\sigma = \frac{\gamma}{\lambda} \text{ dove } \gamma = \frac{C}{C_0} \text{ e } \lambda = \frac{L}{L_0} \quad (2.5)$$

Un valore di  $\sigma$  significativamente più grande di 1 indica la coesistenza dei due attributi della topologia di piccolo mondo, in quanto ci si aspetta  $C \gg C_0$  e  $L \approx L_0$ . È opportuno notare che di per sé il coefficiente  $\sigma$  fornisce solo informazioni sull'architettura globale della rete, mentre non si possono ricavare conclusioni circa l'organizzazione di essa ad una scala intermedia: in particolare, non è possibile dedurre direttamente l'eventuale presenza di moduli.

### 2.2.2 Invarianza di scala: distribuzioni a legge di potenza

Un'altra particolarità emersa dall'analisi delle reti reali è che la distribuzione del grado dei nodi  $P(k)$  mostra un andamento a “legge di potenza”, in contrasto con ciò che era atteso, ossia una distribuzione piccata attorno ad un grado medio come accade nelle reti random e regolari (Fig. 3).

$$P(k) \sim Ak^{-\gamma} \text{ con } 2 < \gamma < 3 \quad (2.6)$$

Le leggi di potenza indicano che una distribuzione del grado dei nodi ad *invarianza di scala*: non esiste una “scala” caratteristica per il grado  $k$  ossia un valore medio

di riferimento attorno al quale  $k$  fluttua con piccoli scostamenti, ma invece esiste un ampio range di variabilità. Ciò è dovuto alla particolare forma funzionale di  $P(k)$ , che resta invariata per riscalamenti della variabile indipendente  $k$ , a meno di un fattore moltiplicativo: in termini qualitativi, si può dire che “zoomando” su un particolare segmento della distribuzione, essa resta della stessa forma. In termini quantitativi, ciò si esprime nel seguente modo:

$$P(ax) \sim A(ax) = Aa^{-\gamma}x^{-\gamma} = Abx^{-\gamma} = bP(x) \quad (2.7)$$

Le reti che presentano questo tipo di distribuzione  $P(k)$  sono dette *reti ad invarianza di scala* [22, 24]. Per esse, la definizione di un grado medio, sebbene matematicamente lecita, ha poco senso: infatti, dal calcolo della varianza attorno tale valore, si trova infatti che essa diverge col secondo momento della distribuzione  $\langle k^2 \rangle$ , formalizzando così l’effetto di una enorme dispersione attorno al valore atteso.

$$\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2 \quad \text{dove} \quad \langle k^2 \rangle = \int_{k \min}^{k \max} k^2 P(k) \sim k_{\max}^{3-\gamma} \quad (2.8)$$

Esiste quindi una gerarchia continua di nodi, in cui coesistono da pochi hubs di alto grado molto connessi ad innumerevoli nodi di basso grado, diversamente dai modelli matematici ideali in cui i nodi sono tutti topologicamente equivalenti, ossia indistinguibili tra loro e con un grado piuttosto omogeneo.

La forma delle funzioni del tipo  $P(k) \sim k^{-\gamma}$  porta naturalmente all’esistenza di lunghe code piuttosto smussate nel limite di grandi  $k$ : nei sistemi finiti, in cui  $k$  è un numero che può variare in un ampio range ma comunque limitato, le distribuzioni a legge di potenza appaiono troncate a partire da un determinato valore di  $k_{\max}$ , a causa della finitezza del numero di nodi  $N$ : questo provoca dei bruschi cut-off o “tagli” alle code, che appaiono “spesse” cioè densamente popolate da nodi, fino a che non si interrompono andando bruscamente a zero. Tale effetto è tanto più evidente quanto più il campione di nodi è limitato oppure se, considerata la totalità dei nodi, esistono limitazioni di tipo spaziale o di bilancio “energetico” allo sviluppo di una rete: ciò

accade soprattutto nelle reti immerse nello spazio euclideo (i.e. reti geografiche) o per tutte quei networks il cui accrescimento richiede un significativo costo in termini di energia, come le reti biologiche. Interessante notare che le reti cerebrali presentano entrambi i tipi di vincoli, essendo immerse in un volume finito, nel caso dell'uomo quello del cervello e del cranio, e richiedendo all'organismo un equilibrio energetico in attivo per garantirne la funzionalità. In questi casi, quando il comportamento ad invarianza di scala esiste per un intervallo limitato di gradi, si parla quindi di larga scala [25].

## 2.3 Modelli ed architetture delle reti reali: modello WS e modello BA

L'esistenza delle proprietà di piccolo mondo e di invarianza di scala nelle reti reali ha stimolato la creazione di modelli adeguati che rispecchiassero tali caratteristiche. In questa sezione, verrà presentato un modello che spiega l'emergere di ciascuna delle due proprietà: la loro importanza sta nel fatto che sono archetipi nel loro genere e di aver aperto la strada alla nuova teoria delle reti: non più la pura astrazione matematica della teoria dei grafi, ma scienza aderente alla realtà sperimentale.

### 2.3.1 Un modello di rete di piccolo mondo: il modello di Watts e Strogatz

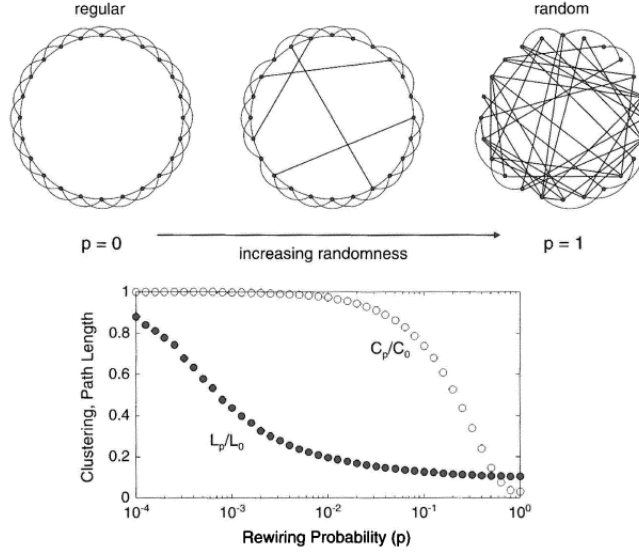
Il modello di Watts e Strogatz [9], in breve *modello WS*, nasce per spiegare l'origine l'effetto di piccolo mondo sulla base dell'analisi degli schemi di connessione tra nodi: è un modello quindi di tipo statico, in cui la struttura non è messa in relazione all'evoluzione e alla crescita della rete stessa. Il modello WS genera grafi  $G_{N,K}^{WS}$  che presentano contemporaneamente sia un breve cammino caratteristico  $L$ , sia un alto coefficiente di clustering  $C$ , segni distintivi della proprietà di piccolo mondo, effettuando una sorta di interpolazione tra un anello regolare ed un grafo random, di pari dimensione  $N$  e stesso numero di archi  $K$ , con la variazione di un solo parametro: la *probabilità di riconnessione*. L'interpolazione avviene partendo da un anello regolare di  $N$  nodi, in cui ciascun nodo è simmetricamente connesso ai suoi  $2m$  nodi vicini, per un totale di  $K = mN$  archi. Per ciascun nodo  $i$ , ogni arco che lo connette ad

un vicino raggiungibile in senso orario da  $i$  viene “ricablato” e connesso ad un altro nodo scelto in modo casuale tra i restanti  $N - 1$  con una *probabilità di riconnessione*  $p$ , mentre viene preservato con una probabilità  $1 - p$ . Di conseguenza:

- se  $p = 0$ , il grafo resta completamente regolare;
- se  $p = 1$ , il grafo diventa completamente random, visto che nessun arco preesistente è conservato, ma viene ricollegato con una probabilità uniforme  $p$ ;
- se  $0 < p < 1$ , il grafo presenta caratteristiche intermedie tra la regolarità e la casualità.

Dallo studio di  $C$  e  $L$  come funzioni di  $p$  [9], si trova che in corrispondenza di valori di  $p$  appena più grandi di 0, il cammino caratteristico  $L$  ha un brusco calo rispetto al dato iniziale mentre  $C$  si mantiene costante intorno al valore (alto) del grafo regolare di partenza: per i valori di  $p$  per cui si verifica questa configurazione, si ottiene una rete di piccolo mondo.

Questo accade perché la riconnessione dei link a nuovi nodi finali genera archi a lungo raggio che connettono link altrimenti distanti, veri e propri shortcuts che accorciano le distanze tra i nodi tipiche dei grafi regolari, lasciando intatta però la densa connettività locale. La rapida diminuzione di  $L$  e la quasi costanza di  $C$  in corrispondenza di un lieve aumento di  $p$  può essere motivato dal fatto che la dipendenza di  $L$  da  $p$  è altamente non lineare: infatti, quando nodi remoti topologicamente vengono connessi, l’abbreviazione del cammino minimo che li lega non è risentita solo dai due nodi coinvolti, ma anche dai nodi ad essi adiacenti. Diversamente, la creazione di questi nuovi link ha un effetto al più lineare sul coefficiente  $C$  che descrive la connettività locale attorno ad un fissato nodo: a seguito di una piccola variazione di  $p$ , esso subirà una diminuzione altrettanto impercettibile.



**Figura 6:** Il modello WS: nel range di probabilità tra  $10^{-2} < p < 10^{-1}$ , si osservano le proprietà di piccolo mondo

### 2.3.2 Un modello di rete ad invarianza di scala: il modello di Barabasi e Albert

Il modello di Barabasi-Albert [22], in breve *modello BA*, spiega la topologia delle reti ad invarianza di scala mettendo in relazione l'esistenza di una distribuzione del grado  $P(k)$  a legge di potenza con il modo in cui il sistema si è evoluto nel tempo: l'idea alla base del modello è che i meccanismi dinamici che hanno guidato lo sviluppo della rete abbiano determinato le sue proprietà topologiche riscontrabili alla fine dell'evoluzione, o comunque ad un tempo successivo rispetto alla sua genesi. Il modello BA, sviluppato inizialmente per analizzare la struttura del World Wide Web, dimostra come la distribuzione  $P(k)$  a legge di potenza sia dovuta ad un particolare processo di crescita, detto “*connessione preferenziale*” (*preferential attachment*), secondo cui i nuovi nodi che si aggiungono alla rete si connettono a quelli preesistenti in maniera proporzionale al loro grado: maggiore è il grado dei nodi preesistenti, maggiore è la probabilità che i nuovi nodi si connettano ad essi e non a nodi di grado inferiore. Questo comportamento è noto anche col nome di “effetto Matteo” ed in sostanza

esprime il concetto, ben noto in economia e sociologia, che “il ricco diventa sempre più ricco”. Nel caso delle reti, la “ricchezza” è data dal grado di un nodo.

In termini quantitativi, secondo il modello BA, un grafo  $G_{N,K}^{BA}$  è costruito a partire da  $m_0$  nodi isolati: a ciascuno step temporale  $t = 1, 2, 3, \dots, N - m_0$ , il cui numero è fissato dalla dimensione finale  $N$  del grafo, un nuovo nodo  $j$  con  $m$  link connessi ad esso è aggiunto al grafo. Ovviamente sarà  $m \leq m_0$ . Nel caso più semplice, la probabilità  $\Pi$  che un arco connetta  $j$  ad un nodo  $i$  già esistente è linearmente proporzionale al grado corrente di  $i$ : in questo caso, si parla di connessione preferenziale lineare.

$$\Pi_{j \rightarrow i} = \frac{k_i}{\sum_l k_l} \quad (2.9)$$

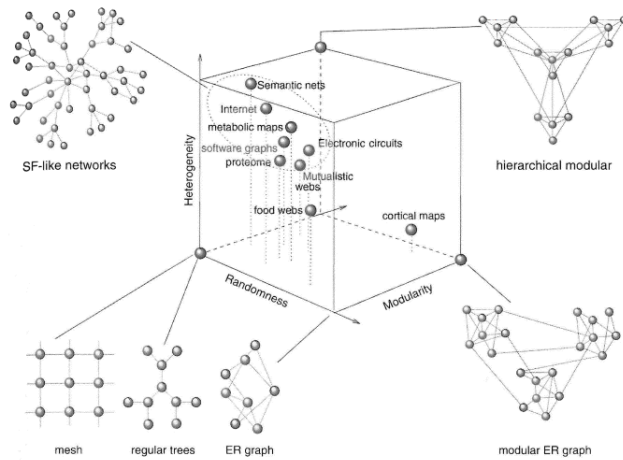
Al denominatore compare la somma totale dei gradi di tutti i nodi come fattore di normalizzazione. Dato che ogni nuovo nodo ha  $m$  archi, la rete ad un tempo  $t$  avrà  $N = m_0 + t$  nodi e  $K = mt$  archi, e la somma dei gradi sarà  $2K$ . In queste ipotesi e nel limite di grandi tempi, il modello produce una distribuzione  $P(k) \propto k^{-\gamma}$  con  $\gamma = 3$ : il valore di  $\gamma = 3$  fissato è uno dei limiti del modello ed ha condotto a successive generalizzazioni, ottenute ad esempio considerando una probabilità di connessione non lineare rispetto al grado dei nodi preesistenti. Tuttavia, per i grafi ottenuti col modello BA originale, è stato stimato [24] che il cammino caratteristico  $L$  cresce logaritmicamente con  $N$  ed è più piccolo rispetto a quello di grafo random di pari  $N$  e  $K$ ; invece, il coefficiente di clustering  $C$  decresce con  $N$  secondo la legge  $C \approx N^{-\frac{3}{4}}$ , più lentamente che in un grafo random dove  $C \approx N^{-1}$  ed ancora diversamente rispetto ad una rete di piccolo mondo in cui  $C$  è praticamente costante rispetto ad  $N$ .

### 2.3.3 Il problema della modularità e la classificazione delle reti

Le quattro architetture più studiate nella scienza delle reti, nonché le più diffuse tra le reti reali sono reti random, regolari, di piccolo mondo e ad invarianza di scala: tuttavia, nessuno di questi modelli prevede un'altra caratteristica osservata nei networks reali, peraltro di fondamentale importanza per le reti cerebrali cioè la modularità. L'esistenza di moduli, ad un primo approccio sembra essere legata strettamente alla proprietà di piccolo mondo: infatti, poiché i nodi all'interno di uno stesso modulo

sono fortemente connessi tra loro, il numero di triangoli in una rete modulare è molto più alto rispetto a quello in un grafo random, portando così ad un elevato coefficiente di clustering. Inoltre, la presenza di pochi archi che connettono nodi appartenenti a moduli differenti crea i cosiddetti shortcut, fondamentali affinché sia abbia un piccolo cammino caratteristico. Analizzando una rete modulare quindi, si possono quindi rilevare le proprietà tipiche di una rete di piccolo mondo: tuttavia non è vero il viceversa cioè non è vero che tutte le reti di piccolo mondo sono modulari. Ad esempio, il modello WS genera reti di piccolo mondo non modulari. Di conseguenza, le tecniche di ricerca della modularità vanno integrate in ciascun modello di rete reale, se si desidera mettere a fuoco tale aspetto. Un altro aspetto rilevante è che la modularità nelle reti reali si presenta a diverse scale di organizzazione: la rete nel suo complesso può essere suddivisa in moduli che a loro volta possono essere suddivisi in sottomoduli e il partizionamento è ricorsivo: in questo caso si parla di *modularità gerarchica*. La rete è organizzata in moduli a diversi livelli, ciascuno strettamente legato al livello superiore ed inferiore attraverso una relazione dimensionale: se la gerarchia dei moduli è esatta, allora può essere definita come “frattale”. In molte reti reali, specialmente quelle biologiche, l’autosimilarità tra i moduli appartenenti ai vari livelli è più una proprietà statistica che esatta e nel caso delle reti cerebrali la struttura modulare è approssimativamente invariante su un numero finito di livelli gerarchici [26].

L’emergere di una struttura modulare nelle reti reali, ed in particolare in quelle che schematizzano sistemi complessi legati all’elaborazione di informazioni e segnale, come d’altronde sono quelle cerebrali, è legata a ragioni funzionali: un sistema scomponibile in componenti poco connessi tra loro ma fortemente connessi al loro interno, consente un veloce adattamento ad un cambiamento delle condizioni ambientali e di conseguenza una più semplice evoluzione del sistema in risposta a tali cambiamenti. Questo è giustificato dal fatto che i sistemi modulari posso evolvere variando o duplicando un solo modulo alla volta, lasciando tutti gli altri invariati dato che a causa delle scarse connessioni intermodulari, essi sono di fatto isolati l’uno dall’altro, adottando un punto di vista locale. Questo isolamento locale consente alla rete di continuare a funzionare anche nel momento in cui un modulo sta subendo



**Figura 7:** Diagramma dell'architettura delle reti [28].

modificazioni, rendendo quindi la rete più robusta [27].

Tenendo conto anche delle proprietà di modularità, le reti reali possono essere classificate ed organizzate in uno “spazio di possibili reti” [28]: oltre agli specifici networks reali, ciascun modello di rete occupa una posizione all'interno di questa scatola tridimensionale in cui tre assi ortogonali tra loro indicano ciascuno una caratteristica ben precisa dei grafi:

- randomness: indica la propensione di un grafo ad avere proprietà simili una rete random ER. A valori nulli di randomness troviamo i grafi regolari, mentre per valori crescenti troviamo appunto i grafi random;
- eterogeneità: relativamente al grado dei nodi, indica quanto questo è omogeneo attorno ad un valor medio o invece varia in un ampio range di valori interi. A valori nulli di eterogeneità troviamo i grafi regolari, ai loro antipodi quelli ad invarianza di scala;
- modularità: indica la tendenza di grafo ad essere organizzato in moduli.

Si nota che le mappe corticali, uno specifico tipo di rete cerebrale, occupano una posizione tale che per esse si può desumere un buon valore di randomness e quindi

brevi cammini caratteristici, un deciso carattere modulare e una non trascurabile eterogeneità del grado dei nodi.

## 3 Connettività del cervello umano: il connettoma

### 3.1 La prospettiva delle reti complesse nello studio del cervello umano

#### 3.1.1 Motivazioni

L'idea che il cervello sia un sistema complesso è ben radicata in biologia e nelle neuroscienze sin dalla metà del diciannovesimo secolo: tale interpretazione emergeva in nuce nella visione di Golgi nel considerare il cervello come un reticolo di neuroni connessi in maniera continua ed organica e non come un insieme di cellule nervose indipendenti ed autonome dal punto di vista funzionale, la cosiddetta “dottrina del neurone”, incapace di spiegare la funzione integrativa del sistema nervoso umano [3]. In questa ottica, l'approccio delle reti complesse nello studio del cervello umano e più in generale del sistema nervoso degli esseri viventi sembra essere una scelta quasi naturale: il cervello è il sistema complesso per eccellenza in cui i suoi componenti elementari, i neuroni, si organizzano tramite le sinapsi in differenti patterns di connettività, anche in modo dinamico nel tempo, in base alle diverse esigenze funzionali. L'elaborazione di risposte complesse, la costruzione di pensieri, la memoria, la reazione ad uno stimolo sensoriale esterno sono tutte attività integrative, ossia che coinvolgono diverse aree del cervello contemporaneamente ed in maniera correlata e sono il frutto di questa azione congiunta, non potendo essere svolte da una singola cellula nervosa. Da questo punto di vista, i metodi della scienza delle reti possono fornire strumenti di calcolo per dimostrare l'effettiva esistenza di questa architettura cooperativa e spiegare come a diversi schemi di connessioni tra neuroni corrispondano aspetti funzionali differenti e ben determinati [3]. Di conseguenza, la “traduzione” del cervello in rete complessa appare un'operazione quasi immediata in linea di principio, sebbene non sia un processo così diretto perché è necessario in primis precisare quale tipo di connettività del cervello si desidera analizzare ed in secondo luogo definire in base alla precedente distinzione cosa rappresenta l'oggetto matematico “rete cerebrale”, cioè quali entità reali siano rappresentate da nodi e archi.

### 3.1.2 Tipi di connettività

Il primo passo per studiare il cervello tramite il mezzo delle reti è definire precisamente i diversi tipi di connettività che si possono analizzare secondo tale strumento matematico. Si possono distinguere tre tipi di connettività, legati a loro volta a precise tecniche di indagine visiva [3]:

1. La *connettività strutturale* descrive le connessioni anatomiche che legano i diversi elementi neurali. Nel caso del cervello umano, per “elementi neurali” si intendono entità differenti a seconda della scala alla quale si analizzano queste connessioni “fisiche”: la connettività strutturale infatti può essere studiata in linea di principio tra singoli neuroni, tra diversi gruppi di neuroni ed a scala più larga tra diverse regioni della corteccia connesse tra loro da fasci di fibre nervose (sostanza bianca). Gli schemi di connettività strutturale sono relativamente stabili per piccoli tempi (dai secondi ai minuti), ma possono essere plastici se osservati su tempi più lunghi (da ore a giorni). Nell’ambito del cervello umano, limitazioni di carattere sperimentale permettono la descrizione di tale connettività per la sola larga scala, come verrà specificato in seguito. Si può pensare alla connettività anatomica o strutturale come una sorta di cinematica delle connessioni cerebrali, nel senso che permette di stabilire e quantificare come due diverse aree dell’encefalo sono collegate e possono scambiare informazioni in maniera più o meno efficiente, ma senza indagare le ragioni di questi collegamenti. Le metodologie sperimentali per lo studio di tale connettività sono le tecniche di tracciamento delle vie neurali; nel caso degli animali, queste possono essere anche invasive, mentre per l’uomo le tecniche d’elezione sono quelle non invasive come la risonanza magnetica strutturale, MRI, o la DTI (Diffusion tensor imaging).
2. La *connettività funzionale* descrive l’accoppiamento dinamico di regioni cerebrali, interpretandolo come segno di una relazione funzionale tra le aree coinvolte. Questo tipo di connettività è studiato tramite registrazioni nel tempo dell’attività neuronale tramite tecniche di elettroencefalografia (EEG) e risonanza magnetica funzionale (fMRI), tramite le quali si ricavano serie temporali

di dati che vengono successivamente analizzati con tecniche statistiche: l'eventuale esistenza di una correlazione tra le diverse zone del cervello è stimata come una deviazione dall'ipotesi di indipendenza statistica tra i segnali generati da tali regioni spesso spazialmente remote. L'emergere di una eventuale correlazione fra essi è sintomo di un legame dinamico, di una dipendenza tra le aree in esame, che perdura solo in un preciso transiente di tempo, spesso brevissimo: la connettività funzionale infatti tende ad evolvere in decimi di millisecondo e ciò rende difficoltoso il suo studio [29]. Bisogna specificare però che data la natura statistica dell'indagine, l'esistenza di patterns di connettività funzionale non implica la presenza di una relazione causale tra l'attivazione dei vari elementi neurali.

3. La *connettività efficace* rappresenta il tentativo di catturare la natura causale delle connessioni tra elementi neurali: è quindi una evoluzione in senso “modelistico” della connettività funzionale che cerca di costruire una rete cerebrale in cui le relazioni fra gli elementi neurali sono di causa-effetto, indotte da stimoli e compiti provenienti dall'esterno. Questo tipo di connettività viene ugualmente studiato tramite registrazioni temporali di segnale, risultando quindi fortemente dipendente dal tempo. Sia la connettività funzionale che quella efficace rappresentano la “dinamica” delle connessioni cerebrali, dal momento che tentano di spiegare le cause e la frequenza statistica con cui determinati patterns di connettività si presentano.

Le tre modalità di connettività non sono esclusive nella descrizione del cervello in termini di rete complessa, nel senso che sono l'una interdipendente dall'altra. Tralasciando la connettività efficace il cui studio è ancora allo stato embrionale, gli aspetti anatomici e funzionali sono strettamente correlati: una rappresentazione completa della connettività strutturale non può non tener conto della fisiologia e delle proprietà biofisiche della conduzione del segnale, caratteristiche che determinano l'andamento temporale delle relazioni tra elementi neurali. Analogamente, gli schemi dinamici di correlazioni statistiche dell'attività elettrica devono necessariamente fondarsi su un substrato di connessioni anatomiche. Questo lavoro di tesi si concentrerà sulla

descrizione dal punto di vista delle reti complesse della connettività strutturale del cervello umano: la mappa completa delle connessioni cerebrali anatomiche è detta *connettoma* ed è uno tra i principali campi di ricerca della fisica applicata alle neuroscienze.

## 3.2 Il connettoma umano

Nonostante l'interdipendenza tra la connettività strutturale e funzionale, la conoscenza del cervello in quanto sistema complesso di parti interagenti non può essere esaustiva in assenza della mappa dei collegamenti fisici fra i suoi elementi strutturali: anche se non si può ridurre lo studio del sistema complesso “encefalo” al solo diagramma dei link anatomici, non c'è dubbio che i collegamenti strutturali plasmano e modellano anche gli schemi di connettività funzionale ed efficace e ne rappresentano il framework anatomico essenziale senza il quale le diverse regioni cerebrali non potrebbero di fatto scambiare informazioni. Di conseguenza, la mappa delle connessioni neurali anatomiche rappresenta lo strumento basilare per poter interpretare la dinamica delle reti cerebrali e consentire l'eventuale formulazione di modelli funzionali: per questo motivo può essere riconosciuto alla connettività strutturale, considerata come “la cinematica”, una sorta di “primato” rispetto alle altre due modalità.

### 3.2.1 Definizione e caratteri fondamentali. L'esempio del C. Elegans

Il *connettoma umano* è “la descrizione complessiva della rete strutturale di elementi e connessioni che formano il cervello umano” [30]. In altre parole, esso è la rete formata dalle connessioni anatomiche fra i diversi elementi neurali del cervello umano, comprensiva delle diverse scale in cui la connettività cerebrale si articola. Le caratteristiche salienti del connettoma emergono dalla sua stessa definizione: in primo luogo, è una rappresentazione dell'organizzazione anatomica dell'encefalo e quindi riguarda l'enorme ma finito insieme di collegamenti fisici tra i vari elementi neurali. Il fatto che sia la descrizione di una realtà materiale funge come una sorta di validazione metodologica, nel senso che metodologie differenti di tracciamento ed imaging devono condurre a risultati convergenti e consistenti tra loro. In secondo luogo, il

connettoma è una *descrizione* della struttura fisica delle connessioni cerebrali: un tale approccio non è finalizzato all'ottenimento di una replica esatta dell'anatomia del cervello in termini di dimensioni e distanze metriche tra i vari elementi neurali, ma invece riassume gli aspetti topologici fondamentali e ne descrive l'organizzazione multiscala, tipica di questo sistema. Di conseguenza, una mappatura simile non riduce la connettività strutturale del cervello al livello delle singole ramificazioni delle  $10^{15}$  sinapsi che connettono i circa  $10^{11}$  neuroni, ma fornisce una rappresentazione globale che spazia attraverso le diverse scale, da quella cellulare a quella delle macroaree della corteccia. La prospettiva di comprendere le diverse scale in un'unica mappa rinunciando al dettaglio fine è anche una esigenza sperimentale: bisogna ricordare che il connettoma è un network non fisicamente accessibile, nel senso che è dedotto dai dati tramite tecniche di imaging e allo stato attuale delle tecnologie disponibili la mappatura delle connessioni anatomiche a livello cellulare non è possibile. In sostanza la strada seguita è una mappatura delle connessioni su larga scala tra aree della corteccia, che naturalmente comprendono in maniera gerarchica quelle delle scale dimensionali inferiori. Il problema del tracciamento rispetto alla dimensionalità verrà affrontato in seguito in modo più preciso. Un ultimo aspetto che emerge dalla definizione di connettoma è che esso è la descrizione di una rete e non solo una grande raccolta di dati: in quanto tale, può essere analizzato tramite gli strumenti della matematica e della scienza delle reti. Grazie a questo legame tra neuroscienze e fisica statistica, il connettoma si configura essere come il fondamento teorico dello studio della funzionalità del cervello, in stretta relazione alla sua struttura [31].

Il nome connettoma è stato scelto in precisa e voluta analogia con il termine genoma, che indica il sequenziamento del DNA di un essere vivente e la sua mappatura: infatti, entrambi descrivono l'azione collettiva e coordinata degli elementi che li compongono sottoforma di rete ed è la struttura stessa a determinarne gli aspetti funzionali. La forte analogia tra i due concetti risiede in molti aspetti: sono entrambe sistemazioni teoriche di dati strutturali che descrivono le interazioni all'interno di sistemi biologici con una dinamica complessa. Entrambi mostrano una variabilità tra organismi di specie differenti, tra quelli della stessa stessa specie e nel corso del tempo anche nell'ambito del singolo individuo. Inoltre sia il connettoma che il ge-

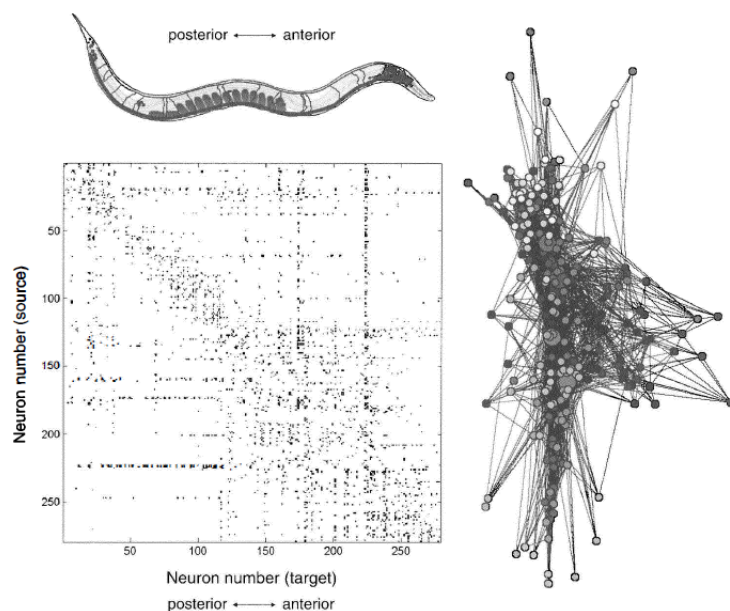
noma sono immersi nello spazio fisico tridimensionale e di conseguenza in entrambi i casi questa relazione pone dei vincoli strutturali al loro sviluppo.

### 3.2.2 Un esempio di connettoma: la mappa neurale del *C. Elegans*

Al giorno d'oggi, l'unico essere vivente di cui si conosce interamente il connettoma è il nematode *Caenorhabditis elegans*, un verme lungo circa un millimetro che vive nel sottosuolo delle regioni dal clima temperato. È questo un caso emblematico poiché la sua mappa completa delle connessioni neurali è nota al livello dei singoli neuroni e sinapsi, dato il loro numero relativamente ridotto rispetto a quello degli organismi superiori: la matrice di adiacenza ottenuta per il connettoma del *C. Elegans* infatti conta 279 nodi (che in questo caso corrispondono ai neuroni) connessi da 6393 sinapsi, 890 giunzioni elettriche e 1410 giunzioni neuromuscolari, per un totale di 8693 archi, tutti numeri molto più piccoli rispetto al caso del cervello umano.

La matrice di adiacenza è stata ottenuta ricostruendo le connessioni neuronali tramite l'analisi di una serie di migliaia di sezioni del verme spesse appena 50 nm analizzate al microscopio elettronico [32]: questa tecnica, assieme alla morfologia semplice del nematode, ha permesso di conoscere in modo preciso la lunghezza dei collegamenti sinaptici e la posizione spaziale di ciascun neurone, consentendo così di studiare la relazione tra la topologia della rete e la sua immersione nello spazio fisico. Analizzando la connettività cellulare dal punto di vista delle reti, i risultati ottenuti indicano una netta abbondanza di caratteristiche non randomiche: le connessioni tra neuroni si stabiliscono in modo predominante tra neuroni prossimi spazialmente, secondo un principio di "economia" nella lunghezza degli archi. Questo aspetto è desumibile osservando l'alta densità di connessioni disposte attorno alla diagonale principale della matrice di adiacenza, costruita disponendo i nodi esattamente nell'ordine spaziale in cui i corrispettivi neuroni sono disposti lungo l'asse longitudinale del verme.

La conseguenza della tendenza a stabilire connessioni tra neuroni vicini spazialmente è la sovrabbondanza di particolari motivi, rispetto alle reti random equivalenti: nel sistema nervoso del *C. Elegans*, prevalgono quelli triangolari, sintomi di una



**Figura 8:** Matrice di connessione e rappresentazione grafica del connettoma del *C. Elegans* [32]

spiccata tendenza all'ottimizzazione della connettività locale piuttosto che di quella globale. Un'ulteriore conseguenza è il carattere di piccolo mondo della rete neurale del *C. Elegans*, come altri studi [9] hanno mostrato, calcolando un coefficiente di clustering  $C$  molto maggiore ed un cammino caratteristico simile (ossia basso) rispetto a rispettivi valori per una rete random equivalente usata come test.

### 3.3 Tecniche sperimentali

Il primo passo verso la codifica del connettoma umano in quanto rete complessa risiede nello studio della connettività strutturale di parti del cervello umano come la corteccia. Tale studio non è una operazione diretta, in quanto il network va dedotto dai dati: fondamentali risultano quindi le tecniche sperimentali di tracciamento e di imaging atte ad estrapolare dati anatomici che andranno tradotti nel linguaggio delle reti, stabilendo in modo univoco cosa rappresentano i nodi e gli archi.

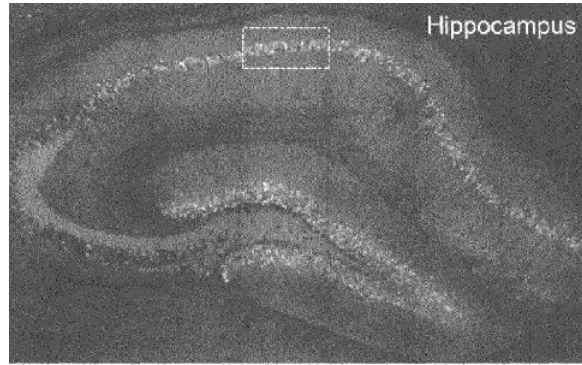
### 3.3.1 La dimensionalità multiscala del cervello umano

La connettività strutturale del cervello umano è una architettura multiscala, organizzata in diversi livelli spaziali annidati l'uno dentro l'altro. Si possono distinguere tre scale di organizzazione:

- la *microscala* o *scala cellulare*, costituita dai singoli neuroni connessi dalle sinapsi;
- la *mesoscala*, costituita dai raggruppamenti anatomici di neuroni e le fibre che li connettono;
- la *macroscala*, formata da regioni corticali distinte collegate da fasci di sostanza bianca ossia fasci di assoni.

A seconda della grana dimensionale alla quale si vuole mappare il connettoma, è richiesta una ben precisa risoluzione spaziale e risulta quindi necessario selezionare i metodi sperimentali più adeguati. Per lo studio della microscala, sono richieste risoluzioni spaziali dell'ordine almeno delle dimensioni cellulari ( $\sim 1\mu\text{m}$ ): le tecniche di imaging cellulare più avanzate utilizzano la micrografia elettronica (EM) e riescono a ricostruire finanche le più piccole ramificazioni dendritiche all'interno di blocchi cubici di tessuto neurale con uno spigolo di centinaia di  $\mu\text{m}$  [33]. Approcci di imaging cellulare ancora più avanzati utilizzano coloranti fluorescenti per contrassegnare singoli neuroni [34] in topi geneticamente modificati, detti “rainbow mouse”: in questi animali, i neuroni e le sinapsi, etichettati da colori, possono essere ricostruiti analizzando le immagini ottenute con un microscopio confocale di sezioni ultrasottili di tessuto. La sovrapposizione in pile di queste sezioni consente di la creazione di un'immagine tridimensionale.

Appare evidente che tecniche *in vivo* di questo tipo non possono essere applicate allo studio del cervello umano: l'utilizzo di ogni tecnica invasiva o di colorazione è escluso, limitando la loro applicazione a campioni post mortem. Gli approcci sperimentali ad oggi esistenti per la mappatura della mesoscala sono essenzialmente tecniche di tracciamento che coinvolgono da una parte l'analisi visiva al microscopio di sezioni istologiche, dall'altra l'iniezione di traccianti in vivo che “colorano” la



**Figura 9:** Ippocampo di un “rainbow mouse”: neuroni colorati con markers fluorescenti [34].

membrana cellulare dei neuroni connessi fra loro. Anche queste tecniche, per la maggior parte invasive ed utilizzate ad esempio nel caso di reti cerebrali di mammiferi come il gatto ed il macaco, non sono utilizzate per studi nell’uomo. Lo studio della larga scala è invece effettuato tramite tecniche di imaging non invasivo come la risonanza magnetica strutturale (MRI) e l’imaging di diffusione (DTI e DSI), risultando di conseguenza le più adatte allo studio del cervello umano. Altre tecniche di studio della larga scala sono l’analisi di campioni post mortem, che incontrano parecchi ostacoli a causa del rapido deterioramento dei tessuti.

Emergono da questa breve disamina le ragioni per le quali la connettività strutturale umana è studiata principalmente rispetto alla larga scala. Al di là della scontata impossibilità oggettiva nell’uso di tecniche invasive in vivo, esistono in primis vincoli di carattere sperimentale: le tecniche più adeguate per ricostruire il connettoma con lo scopo principale di fornire una descrizione “complessiva” ed integrativa del cervello umano come una rete sono quelle di imaging applicate alla larga scala. In secondo luogo, emergono limitazioni di carattere computazionale e strumentale che impediscono ad esempio l’applicazione di metodi di imaging cellulare per il tracciamento degli assoni in campioni post mortem: infatti l’utilizzo di tecniche di grande risoluzione come la micrografia elettronica richiederebbe risorse computazionali per l’acquisizione, storage e manipolazione di una enorme mole di dati, al di là delle possibilità tecnologiche attuali. È stato stimato che un singolo connettoma uma-

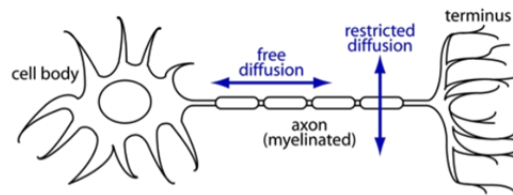
no mappato a livello della microscala comprenderebbe un trilione di gigabytes di dati ossia  $10^{18}$  byte = 1 exabyte. Oltre ai limiti tecnici precedentemente esposti per lo studio del connettoma a livello cellulare, l'analisi micrografica di campioni post mortem non permette di stabilire il collegamento tra l'anatomia e le funzione del campione esaminato, uno tra i principali obiettivi dello studio del connettoma. Infatti, i diagrammi di connessione così ottenuti infatti sono delle "istantanee" congelate nel tempo della connettività strutturale che invece a scala cellulare è plastica ed in continua evoluzione in tempi brevi, dell'ordine dei secondi [31].

### 3.3.2 Tecniche di imaging per la connettività su larga scala

Le tecniche sperimentali più utilizzate per lo studio della connettività strutturale del cervello umano a larga scala sono sostanzialmente tecniche di imaging non invasivo, basate sul principio fisico di risonanza magnetica (RM). Si possono distinguere due approcci distinti: l'Imaging a risonanza magnetica strutturale (MRI) e l'Imaging con tensore di diffusione (DTI e DSI).

L'MRI strutturale si basa sulla classificazione del tessuto cerebrale in base alle differenze del segnale MR prodotto da esse: con questa tecnica è possibile non solo visualizzare come sono disposte spazialmente le diverse aree del cervello, ma anche mostrare le variazioni specifiche delle dimensioni del volume o della superficie di determinate regioni cerebrali. Questo tipo di imaging ha permesso di stimare lo spessore della materia grigia della corteccia, evidenziandone le variazioni tra le diverse aree: è stato dimostrato che l'eventuale correlazione tra lo spessore di due diverse aree della corteccia è associata alla presenza di fibre che legano le due regioni. Su questa tecnica si basa il primo studio sulla connettività strutturale umana a larga scala di He et al, nel 2007, in cui è stata ottenuta la prima matrice di adiacenza per l'intera corteccia [35].

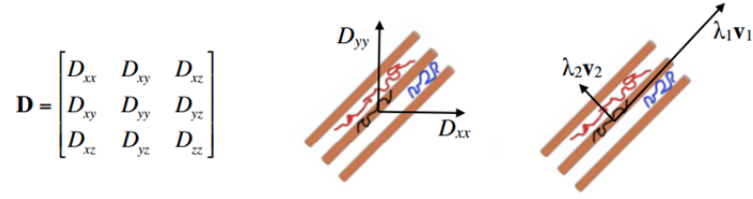
L'Imaging con tensore di diffusione o *Tensor diffusion imaging* è forse la tecnica più usata nello studio delle connessioni tra le aree cerebrali: è una tecnica sempre basata sulla risonanza magnetica che permette di misurare la diffusione delle molecole d'acqua nei tessuti biologici. Nella materia grigia, a variazioni del gradiente



**Figura 10:** Diffusione negli assoni

del campo magnetico di RM, non corrispondono variazioni nella direzione del moto delle molecole di acqua: questo implica che non esistono direzioni privilegiate nella diffusione e di conseguenza la materia grigia è un mezzo isotropo. Invece, negli assoni (Fig. 10) che contribuiscono a formare la sostanza bianca, ossia quel complesso di fibre nervose che connette le diverse aree corticali, la diffusione delle molecole di acqua è fortemente anisotropa: dato che è improbabile che una molecola di acqua attraversi il rivestimento mielinico delle fibre nervose, di conseguenza la molecola si muoverà principalmente lungo l'asse della fibra neurale.

Quindi osservando l'esistenza di una direzione preferenziale di diffusione delle molecole d'acqua, possiamo assumere che la maggior parte delle fibre nella regione in analisi sono orientate parallelamente a questa direzione. La DTI permette di misurare la diffusione nelle tre direzioni spaziali e l'anisotropia nei patterns di diffusione all'interno di porzioni di volume cerebrale detti voxel, una sorta di pixel tridimensionale, associando a ciascuno di essi un tensore tridimensionale, ossia una matrice simmetrica positiva  $3 \times 3$ . Gli elementi di tale matrice sono ottenuti applicando una serie di gradienti di diffusione (cioè variazioni del campo magnetico nel magnete della RM) in maniera che possano fornire almeno 3 vettori indicanti la direzione delle molecole di acqua all'interno di ciascun voxel. L'utilizzo di 6 diversi gradienti è il minimo indispensabile per definire un tensore tridimensionale simmetrico, rappresentato da una matrice con 6 parametri indipendenti. Il tensore quindi descrive la forma tridimensionale della diffusione delle molecole d'acqua e la direzione delle fibre viene indicata dall'autovettore principale del tensore, cioè quello associato all'autovalore di modulo massimo. In sostanza, il tensore viene diagonalizzato e si passa al sistema di riferimento individuato dai suoi autovettori. Per il

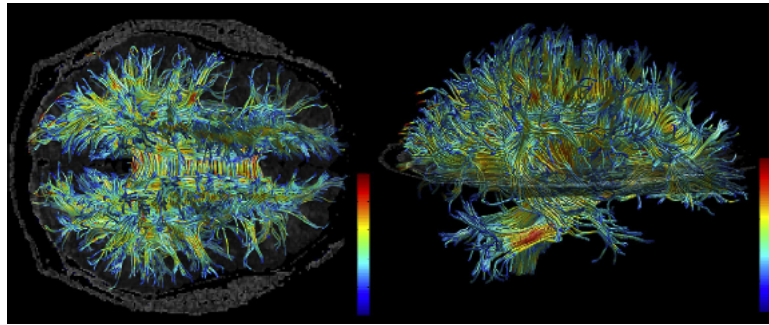


**Figura 11:** Diagonalizzazione del tensore di diffusione

caso più semplice dove la diffusione è isotropica, il tensore di diffusione equivale alla matrice identità moltiplicata per una costante di diffusione  $D$  [3].

Le anisotropie di diffusione vengono poi tradotte in vere e proprie mappe di fibre nervose tramite algoritmi di trattografia che tentano di ricostruire graficamente l'andamento delle fibre di sostanza bianca in maniera consistente con i dati della DTI. Questi algoritmi possono essere distinti in due categorie differenti a seconda della filosofia con cui operano: gli algoritmi deterministici, che cercano di determinare le linee di flusso ottimali del campo tensoriale di diffusione, mentre gli algoritmi probabilistici tentano di stimare tramite mezzi statistici l'esistenza di una fibra, non dandone per certa l'esistenza. Un primo limite della DTI è che non riesce a stabilire il verso dell'anisotropia della diffusione delle molecole di acqua: questo ha come conseguenza che i grafi estratti dai dati con questa tecnica sono tutti non orientati. Un altro limite della DTI è che riesce ad estrapolare una sola direzione di diffusione per voxel, rendendo impossibile l'utilizzo della tecnica per l'analisi di quei voxel in cui si incrociano diverse fibre. Per ovviare a questo limite, si utilizza una variante della DTI detta DSI ossia *Diffusion spectrum imaging*, capace di distinguere direzioni eterogenee di fibre nervose all'interno dello stesso voxel [36].

La validazione delle tecniche di imaging ottenuta tramite la comparazione coi risultati del tracciamento anatomico in vivo è impossibile per l'uomo: gli unici dati da tracciamento anatomico sono quelli post mortem e sono piuttosto limitati. È tuttavia possibile ricavare una loro validazione indiretta nel caso dell'uomo, facendo riferimento alle comparazioni effettuate per altre specie di esseri viventi. È stato dimostrato in numerosi studi effettuati sulla corteccia del macaco un significativo grado di sovrapposizione tra i risultati ottenuti con la DTI e quelli ottenuti tramite



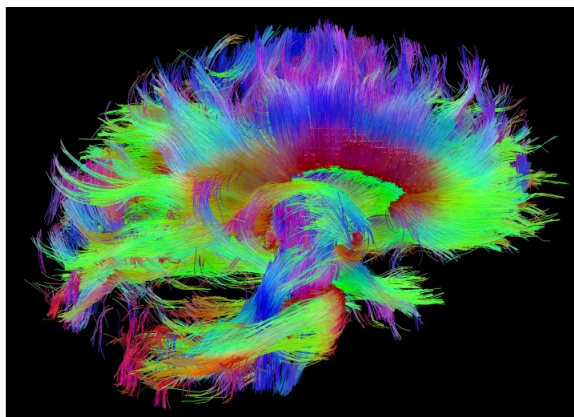
**Figura 12:** Risultato da DTI in paziente schizofrenico. Il codice colore individua l'anisotropia del moto delle molecole di acqua. Il rosso indica alta anisotropia; il blu indica isotropia. Da Kubicki et al., 2007, J. Psychiatric Research, 41.

il tracciatura in vivo, applicati allo stesso singolo esemplare [3]; inoltre, ulteriori studi tramite DSI effettuati su un emisfero cerebrale di un singolo esemplare di macaco, hanno permesso di ricavare una matrice di adiacenza in cui solo il 6% delle connessioni stimate non coincide con quelle ricavate dal tracciamento con i markers. Un completo accordo tra le due tecniche potrebbe non essere mai raggiunto in quanto entrambe operano in due differenti scale specifiche e con risoluzioni differenti: di conseguenza, alcune connessioni possono essere invisibili per una delle due tecniche. Nonostante questi limiti, le tecniche di imaging costituiscono il mezzo migliore per l'analisi della connettività dell'intero cervello a larga scala poiché sono in grado di cogliere una grande mole di dati in un tempo limitato e con una qualità di acquisizione omogenea.

### 3.3.3 Procedura di inferenza del network strutturale dai dati

Una volta fissata la tecnica di imaging per ottenere i dati, va stabilita una procedura che permetta di costruire partendo da essi il network strutturale, sul quale andranno effettuate le misure delle proprietà tipiche delle reti complesse [3, 31, 36]. I passi da seguire sono i seguenti:

1. definizione dei nodi: nel caso delle reti strutturali di larga scala, essi sono identificati con regioni anatomiche definite tramite dati istologici o dati ottenuti tramite le tecniche di imaging;



**Figura 13:** Trattografia di fibre di sostanza bianca. I colori indicano la direzione delle fibre. Rosso=destra-sinistra; verde=anteriore-posteriore; blu=ascendente-discendente. Riprodotta da [www.humanconnectomeproject.org](http://www.humanconnectomeproject.org)

2. stima di una regola di associazione tra coppie di nodi, che consenta la definizione dei links;
3. costruzione della matrice di associazioni tra nodi, che viene raffinata rimuovendo le interazioni “deboli” definendo un valore di soglia rispetto alla regola di associazione tra nodi sopra definita, in modo da ottenere una matrice di adiacenza che rappresenta un grafo binario;
4. calcolo dei parametri tipici delle reti per il network rappresentato dalla matrice di adiacenza ottenuta al punto precedente;
5. comparazione dei parametri calcolati con quelli ottenuti per un insieme di reti random di pari dimensioni e numero di archi, usate come test per evidenziare eventuali caratteri non-random.

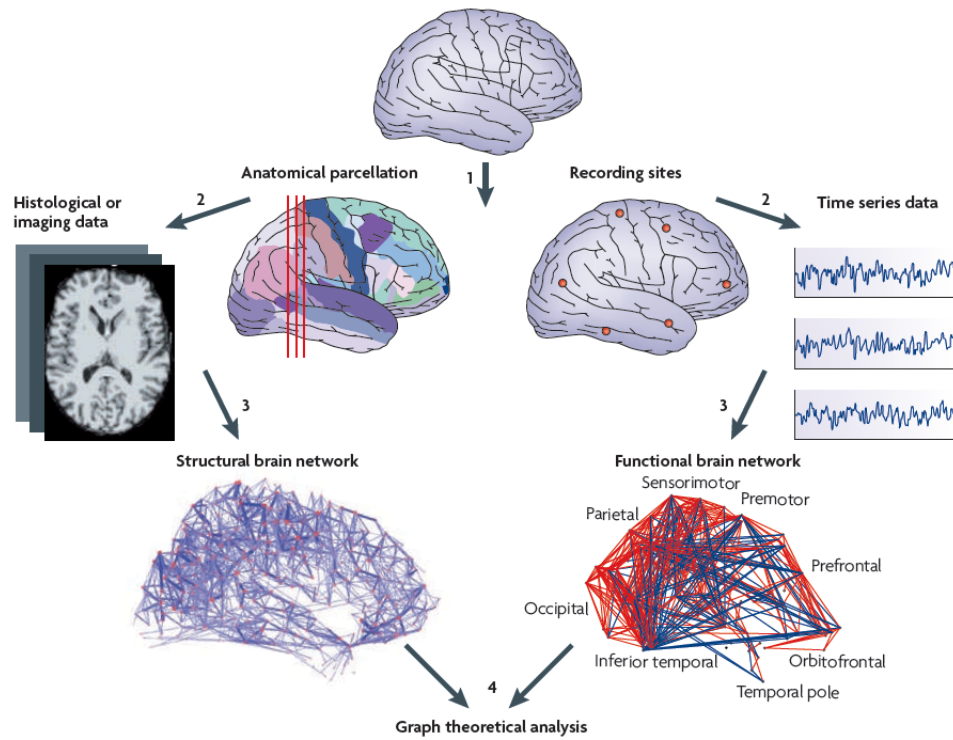
La procedura consta di pochi passi, ma ciascuno di essi mostra problemi metodologici ed cruciale per il risultato finale nel senso che scelte modellistiche errate possono condurre alla produzione di reti che sono descrizioni falsate del sistema complesso che si desidera studiare.

Il problema principale risiede nel primo step: la definizione dei nodi nelle reti cerebrali non è una scelta banale. Di primo acchito, la scelta più naturale sembra

essere quella di identificare i nodi coi neuroni e di conseguenza gli archi con le sinapsi, ma è evidente che questa attribuzione non è compatibile con la mappatura di larga scala che il connettoma si prefigge di fare: infatti, le tecniche di imaging a larga scala registrano dati relativi a larghe porzioni della cervello. Di conseguenza la definizione dei nodi in tutti gli studi sulla connettività strutturale si basa sulla partizione in regioni coerenti del volume cerebrale. Esistono differenti schemi di partizionamento: i più semplici sono basati sui confini anatomici delle aree cerebrali e risultano fin troppo imprecisi. Un approccio alternativo più efficiente consiste nel considerare come nodi raggruppamenti coerenti di voxel in cui è diviso l'intero volume cerebrale: il raggruppamento dei voxel deve dare origini a partizioni di opportuna fittezza. Una partizione troppo rozza genererà perdita di informazione circa alcune connessioni, mentre una troppo fine potrebbe contenere molte informazioni ridondanti, dal momento che molti nodi sono copie di altri. La strada intrapresa è il partizionamento random della corteccia in regioni di dimensioni uguali, che definiscono i nodi: questa scelta è un buon compromesso poiché permette di cercare la partizione migliore variando la scala della suddivisione.

Anche il secondo passo, la definizione degli archi è problematico in quanto diverse “regole” di associazione fra nodi possono essere scelte: nel caso dei networks strutturali dedotti dalla DTI e dalla DSI, la selezione della regola appare semplificata in quanto due aree cerebrali risultano connesse se esse appaiono legate da fibre di sostanza bianca ricostruite tramite imaging.

Fondamentale risulta anche il ruolo della “soglia” rispetto alla quale sono scartate le interazioni “deboli” fra due nodi: differenti soglie determinano matrici di adiacenza più o meno popolate e di conseguenza grafi più o meno connessi. Per questo motivo, i parametri delle reti sono computati per grafi ottenuti con varie regolazioni della soglia. In caso di networks dedotti tramite DTI, l'intervento della soglia va ad agire sulla ricostruzione delle fibre tramite trattografia. Ad esempio, se è stato usato un algoritmo di ricostruzione di tipo probabilistico, viene fissata una data probabilità di riferimento: verranno rigettate le ricostruzioni di fibre (ossia i link) che esistono con un valore di probabilità minore del riferimento mentre verranno tenute solo quelle caratterizzate da valori più alti. L'utilizzo di una soglia è necessario solo se si



**Figura 14:** Procedura di inferenza del network strutturale dai dati di imaging. A sinistra, il processo illustrato nel paragrafo 3.3.3. A destra, l'analogo per la connettività funzionale [36].

desidera ottenere un grafo binario: l'utilizzo della teoria dei grafi pesati permette di oltrepassare questo step.

## 4 Architettura delle reti cerebrali strutturali della corteccia cerebrale umana

### 4.1 I principali studi sulla connettività cerebrale umana a larga scala

I principali studi condotti con l'approccio delle reti complesse che hanno permesso di ricavare una matrice di adiacenza per l'intero cervello umano sono 4 e si differenziano per la tecniche di imaging e per gli schemi di partizionamento impiegati.

Il primo studio effettuato da He et al. [35] nel 2007 studia la connettività su larga scala della corteccia umana dividendola in 54 porzioni, equivalenti ai nodi, e per ciascuna di esse è stata misurato lo spessore corticale tramite tecniche MRI in vivo. Due aree si considerano connesse, e quindi un link è stabilito tra due nodi, se presentano una correlazione significativa tra rispettivi valore dello spessore corticale: il parametro dello spessore corticale è stata scelto come “regola” di connessione in seguito di studi biologici che hanno mostrato che l'esistenza di una correlazione del valore di tale parametro tra due regioni differenti è indice di processi metabolici e trofici simili tra esse, implicando così una stretta correlazione strutturale [31]. Lo studio delle correlazioni è stato effettuato su un campione di 124 soggetti sani, appartenenti al Consortium for Brain Mapping (ICBM), un database di dati ottenuti da imaging MRI, computando per ciascuna coppia di valori di spessore delle 54 aree cerebrali il coefficiente di Pearson, ottenendo così una matrice di correlazione. Eliminando le correlazioni deboli tramite la scelta di una soglia, è stata quindi costruita la matrice di adiacenza, in cui delle possibili  $54 \times 53/2 = 1431$  connessioni, dopo l'applicazione della soglia ne sono rimaste solo 104 significative, cioè solo il 7,3%. Essa è simmetrica e binaria: ha come elementi 1 se la correlazione statistica tra due regioni è significativa, altrimenti 0 se non lo è. Questa matrice binaria è la rappresentazione matriciale del network, sulla quale vengono effettuati le stime dei parametri caratteristici quali clustering C (Eq. 1.8) e cammino caratteristico L (Eq. 1.13).

Altri studi invece utilizzano come tecniche di imaging la DTI e la DSI: in particolare si distingue lo studio di Gong et al.(2009) [37] in cui è stato adottato un

schema di partizionamento in 78 regioni cerebrali per ricostruire una rete di connettività strutturale media grazie ai dati raccolti tra 80 giovani partecipanti, tramite DTI. Ciascuna delle 78 regioni rappresenta un nodo: due regioni sono considerate connesse se l'algoritmo di trattografia prevede con una probabilità significativa che esse rappresentano due punti terminali tra un fascio di fibre ricostruito, ossia se la probabilità che il fascio di fibre ricostruito prosegua verso altre regioni contigue a quelle in esame è molto bassa. Grazie all'utilizzo di criteri statistici che fungono da soglia, dal massimo di  $78 \times 77/2 = 3003$  possibili connessioni tra coppie di nodi, le uniche significative risultano essere solo 329.

Miliari anche i due studi di Hagmann et al (2007,2008) nei quali è stata adottata la DSI: la particolarità dei due studi sta nel fatto che entrambi usano una parcellazione più fine che consente di ottenere una più elevata risoluzione rispetto agli studi precedenti. Inoltre, un'altra fondamentale differenza rispetto agli studi precedenti sta nel fatto che si sono ottenuti networks per ogni singolo soggetto in esame e non più networks "medi": ciò ha consentito il confronto tra i diversi soggetti, aprendo nuove strade per un utilizzo in senso diagnostico della tecnica.

Nel primo studio [38], la corteccia cerebrale è stata modellizzata in primo luogo in voxels, interpretabili come unità elementari dell'immagine 3D ottenuta tramite la DSI e la trattografia; successivamente, i diversi voxels sono stati raggruppati in ROIs ossia *regioni di interesse* in cui l'intera superficie è stata divisa. La dimensione del singolo ROI è un parametro che si fa variare, per il quale esiste un limite inferiore di 1 voxel per ROI. Si preferisce che a ciascun ROI corrispondano più voxels per fare in modo che ciascuna regione sia connessa al resto agli altri ROIs con un numero statisticamente significativo di fibre, tale da giustificare la ricostruzione trattografica. Nella traduzione in network, ciascun ROI corrisponde ad un nodo: il numero di essi varia a seconda della dimensione del ROI, andando da un minimo di 500 ROI con 64 voxel per ROI ad un massimo di 4000 con 8 voxel per ROI. Disponendo di un campione di 2 individui volontari, si è deliberatamente scelto di studiare le proprietà dei network su reti di circa  $N \approx 1000$  nodi e  $K \approx 50000$  archi (per il volontario 1:  $N_1 = 1013$  e  $K_1 = 47217$ ; per il volontario 2:  $N_2 = 956$  e  $K_2 = 50199$ ), in quanto è stato provato che le proprietà calcolate per le altre granularità comprese tra 500 e

4000 davano risultati tutti molto simili fra loro.

Il secondo studio di Hangmann et. Al (2008) [39] si basa sulle stesse tecniche ossia DSI e suddivisione della corteccia in ROIs. La sua peculiarità sta nel modo in cui i 998 ROIs sono stati ottenuti suddividendo a loro volta 66 regioni corticali più vaste ottenute seguendo i confini anatomici della corteccia, su un campione di 5 partecipanti.

Una comparazione diretta tra le matrici di adiacenza ottenute nei diversi studi è difficile poiché si utilizzano differenti schemi di parcellazione, che definiscono a loro volta nodi e archi, mentre una comparazione consistente può essere fatta tra i parametri ottenuti da esse. La mancanza di standardizzazione nelle procedure di imaging, di trattografia e di elaborazione dei dati è un limite di questo campo scientifico in crescente espansione. Per questa ragione, non è ovvio stabilire quale degli approcci citati sia il migliore nel rappresentare le proprietà di connettività strutturale del cervello umano e solo un punto di vista onnicomprensivo è al momento la strategia migliore per affrontare il problema: ciascuna tecnica infatti, con la sua specifica risoluzione spaziale, riesce ad evidenziare patterns di connessioni a livelli diversi [3].

## **4.2 Le reti cerebrali strutturali non sono random**

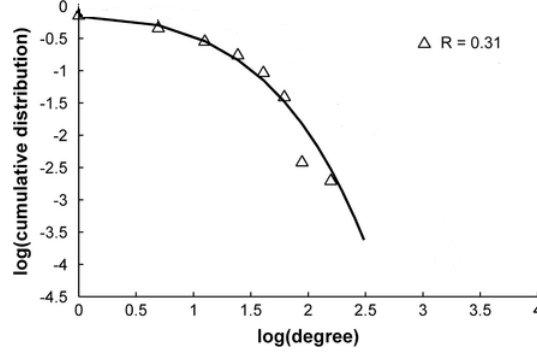
### **4.2.1 Proprietà di piccolo mondo**

Da tutti gli studi sulla connettività su larga scala del cervello umano, emergono attributi non random per i networks cerebrali: analogamente, ciò è riscontrabile anche nei lavori di studio dell'anatomia di larga scala della corteccia dei mammiferi come il gatto ed il macaco [40,41]. Più precisamente, tutti gli studi precedentemente citati sulla connettività strutturale della corteccia umana confermano la presenza dell'architettura di piccolo mondo, dal momento che si ritrovano i due segni tipici che individuano questo tipo di struttura: un alto coefficiente di clustering  $C$  (Eq. 1.8) e un cammino caratteristico  $L$  (Eq. 1.13) simile, se rapportati ai corrispettivi parametri valutati per un insieme di grafi random di stesse dimensioni, i.e. con stesso numero di nodi ed archi e stessa distribuzione del grado  $P(k)$ . Ricordiamo infatti

che una rete di piccolo mondo soddisfa le seguenti condizioni:  $\gamma = C_p^{real}/C_p^{rand} > 1$  e  $\lambda = L_p^{real}/L_p^{rand} \sim 1$ , dove il pedice  $p$  indica che tali parametri mostrano una dipendenza dal valore della soglia utilizzata per ottenere la matrice di adiacenza. Le due condizioni precedenti posso essere rappresentate in maniera sintetica dal coefficiente di piccolo mondo (Eq. 2.5)  $\sigma = \frac{\gamma}{\lambda}$ . Questo implica infatti che il rapporto tra  $C_p^{real}$  e  $C_p^{rand}$  è più grande in valore assoluto rispetto a quello tra  $L_p^{real}$  e  $L_p^{rand}$ , che tende ad 1.

Nello studio [35], che ottiene il network cerebrale tramite l'analisi della correlazione tra lo spessore della corteccia cerebrale delle 54 aree in cui essa è stata partizionata, computando i valori di  $C_p^{real}$  e  $L_p^{real}$  in corrispondenza del valore di  $p$  per il quale si ottengono solo le 104 connessioni significative, si ottiene  $C_p^{real} = 0.30$  e  $L_p^{real} = 3.05$ . Confrontando tali risultati con i valori  $C_p^{rand} = 0.13$ ,  $L_p^{rand} = 2.65$  calcolati e mediati per 1000 reti random equivalenti, si riconosce chiaramente la proprietà di piccolo mondo in quanto  $\gamma = 2.36$ , mentre  $\lambda = 1.15$ . Segue che  $\sigma = 2.04$ , un valore sensibilmente diverso da 1. Nello studio, è stato affrontato anche il problema della sensibilità di tali parametri al variare della soglia: per piccoli valori della soglia, quando tutte le connessioni possibili sono accettate come esistenti, la corrispondente rete cerebrale è praticamente casuale e quasi non distinguibile dalle reti random ad essa abbinate. Al crescere della soglia di correlazione invece, aumenta il numero relativo di connessioni locali e cammini più brevi rispetto a quelli di lungo raggio che vengono invece preferenzialmente scartati: di conseguenza il coefficiente di clustering di aumenta rapidamente, mentre il cammino caratteristico resta quasi invariato: ciò conduce ad un aumento del valore dell'indice di small world  $\sigma$ . Emerge il fatto quindi che l'architettura di piccolo mondo della connettività strutturale umana è legata all'esistenza di una grande interconnessione locale delle varie regioni: in altre parole, due regioni topologicamente vicine risultano anche contigue spazialmente, promuovendo così l'efficienza locale nello scambio delle informazioni. Un'altra misura eseguita sul network ottenuto in [35] è stata la distribuzione del grado: la  $P(k)$  ottenuta (Fig. 15) non può essere una legge di potenza poiché non risulta perfettamente lineare nell'ambito di un grafico bilogarithmico.

La distribuzione è invece meglio rappresentata da una legge di potenza troncata



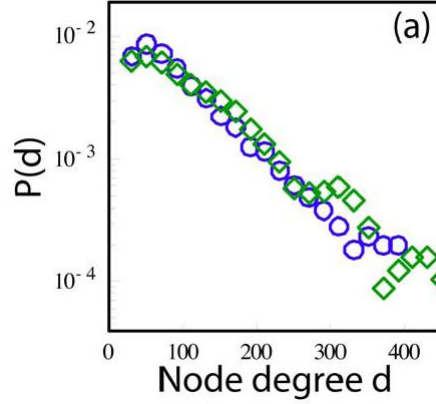
**Figura 15:** Distribuzione  $P(k)$  per il valore più stringente della soglia ottenuta per la rete studiata in [35]

esponenzialmente  $P(k) \sim k^{\alpha-1} e^{-\frac{k}{k_c}}$  con un esponente  $\alpha = 1.38$  e un grado di cutoff  $k_c = 2.50$  in corrispondenza dal valore di soglia più stringente. Questo modello indica che la rete cerebrale ottenuta non è ad invarianza di scala, ma piuttosto a larga scala: sono presenti alcuni hubs con molte connessioni ma è altamente improbabile che vi siano hubs di grado più elevato di 3.

Risultati analoghi sono ricavati dal lavoro di Gong et al [37] in cui emergono nuovamente caratteristiche di piccolo mondo. In più, è stata rivelata l'esistenza di parecchie regioni ad alta centralità, rappresentate nel contesto della rete ottenuta da nodi ad elevata betweenness (Eq. 1.17). Infatti, in corrispondenza del valore di soglia più stringente per cui le connessioni significative risultano essere 329 sulle 3003 possibili, è stato calcolato che il coefficiente di clustering  $C_p^{real} = 0.49$  che risulta essere 4 volte maggiore rispetto a quello  $C_p^{rand} = 0.12$  ottenuto per una rete random equivalente, mentre invece il cammino caratteristico  $L_p^{real} = 2.32$  è quasi del tutto equivalente al corrispettivo valore random  $L_p^{rand} = 2.02$ . Di conseguenza, segue che  $\gamma = 4.07$ ,  $\lambda = 1.15$  e  $\sigma = 3.54$ : risalta pienamente la proprietà di piccolo mondo. Anche in questo studio, è stata evidenziata la dipendenza delle misure di piccolo mondo dal valore della soglia  $p$ , riscontrando lo stesso andamento dei valori di  $C$  ed  $L$  al variare di  $p$ , già analizzato in [35]: gli attributi di piccolo mondo tendono ad enfatizzarsi nel momento in cui la soglia viene resa più stringente.

La valutazione della betweenness per ciascuno dei 78 nodi-regioni e dei 329 links,





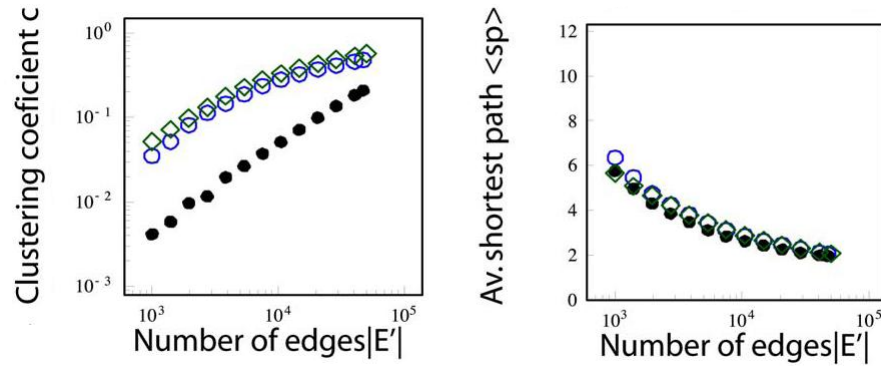
**Figura 17:** Distribuzioni  $P(k)$  per i grafi ottenuti in [38]

dio e distribuzione  $P(k)$ . Nel primo studio, ciascuna misura è stata valutata per entrambi gli individui in maniera distinta: nei grafici, il pallino vuoto rappresenta il soggetto 1 ( $N_1 = 1013$ ), mentre il rombo rappresenta il soggetto 2 ( $N_2 = 956$ ). La distribuzione  $P(k)$ , rappresentata in un grafico con scala log-lin (Fig. 17), è approssimativamente lineare, ad eccezione della coda: ricompare quindi l'andamento a larga scala, rappresentato da una legge di potenza troncata da una coda esponenziale, già postulato negli studi precedenti.

Tuttavia, a differenza degli studi precedenti in cui le reti estrapolate dai dati erano troppo piccole per dedurre con confidenza statistica un preciso andamento funzionale, in questo caso la deduzione è molto più sensata.

Il coefficiente di clustering e il cammino caratteristico sono stati valutati, come da procedura, rispetto ai corrispettivi valori medi di tali parametri per 10 grafi random di 1000 nodi. È stato ricavato l'andamento di  $C$  ed  $L$  e delle rispettive versioni random (Fig. 18), al variare del numero di archi tra un minimo di 1000 fino ad un massimo rispettivamente di  $K_1$  e  $K_2$ , valori rispetto ai quali si trovano i valori di  $C$  ed  $L$  propri delle reti in esame. I sottoinsiemi più piccoli di archi in corrispondenza dei quali sono stati calcolati i coefficienti sono stati costruiti in modo casuale scegliendo tra le due popolazioni di archi disponibili.

Il grafico relativo a  $C$  è di tipo bilogarithmico, mentre quello relativo ad  $L$  è con



**Figura 18:** Clustering e cammino caratteristico in [38]

scala log-lin. Le misure relative al set di grafi random equivalenti sono rappresentate da un pallino annerito. Si nota che per ogni numero  $E'$  di archi nel range considerato, i grafi reali mostrano un coefficiente di clustering molto maggiore rispetto alla rete random di riferimento mentre simili appaiono i cammini caratteristici: tali rapporti tra i coefficienti, sintomo netto di proprietà di small world, si mantengono al variare del numero di archi. Questo metodo è un modo indiretto per dimostrare che la proprietà di piccolo mondo è indipendente dalla granularità della partizione ossia della sua risoluzione (variare il numero di archi infatti equivale a variare le dimensioni del network, dato che alcuni nodi non risulteranno più connessi), ma piuttosto è una caratteristica intrinseca della connettività cerebrale, che risponde ad esigenze di tipo funzionale.

Il secondo lavoro di Hagmann et al. [39], oltre a riconfermare i risultati ottenuti con lo studio precedente rispetto al quale è molto simile, presenta il calcolo dell'assortatività, una quantificazione dei motivi triangolari ed il computo preciso del coefficiente di piccolo mondo  $\sigma$  per il networks i cui nodi rappresentano i 998 ROIs in cui la corteccia è stata divisa. I parametri ottenuti sono valori medi, estratti dal campione dei 5 individui partecipanti. L'indice di small world medio calcolato è  $\sigma = 10.64 \pm 0.61$ , ottenuto confrontando i valori medi di  $C$  ed  $L$  per il network reale con i rispettivi valori medi ottenuti da una serie di 100 grafi random equivalenti. Il valore dell'indice  $\sigma$  molto elevato è indicativo di una fortissima tendenza della re-

te verso una organizzazione di piccolo mondo ed ad un grande valore di clustering: ciò si riflette ad esempio nell'abbondanza di motivi all'interno dei cluster tra nodi. Valutando l'occorrenza statistica  $Z$  (Eq. 1.11) per i motivi triangolari rispetto alla serie di 100 reti random di controllo, ottiene  $Z > 500$ , preciso segnale di deviazione da un'architettura di tipo casuale.

Inoltre, è stata trovata assortatività (Eq. 1.6) media positiva pari a  $0.1941 \pm 0.0513$ : questo indica che i ciascun nodo in ciascuna delle 5 reti esaminate tende ad essere connesso a nodi di grado simile. In particolare, ciò implica che ROIs con alto grado si collegano preferenzialmente fra loro, generando un nucleo strutturale di regioni del cervello (insiemi di hubs) strettamente connesse fra loro tramite hubs: dal punto di vista delle reti, tali insiemi di hubs e nodi a loro volta densamente connessi sono dei moduli. Utile è il confronto con l'assortatività media valutata per 100 reti random di controllo, equivalenti per numero di nodi e loro grado: per esse, si ottiene un valore medio di assortatività pari a  $-0.0122 \pm 0.0098$ , il che implica che i singoli valori di assortatività sono strettamente distribuiti attorno allo zero come atteso per grafi random.

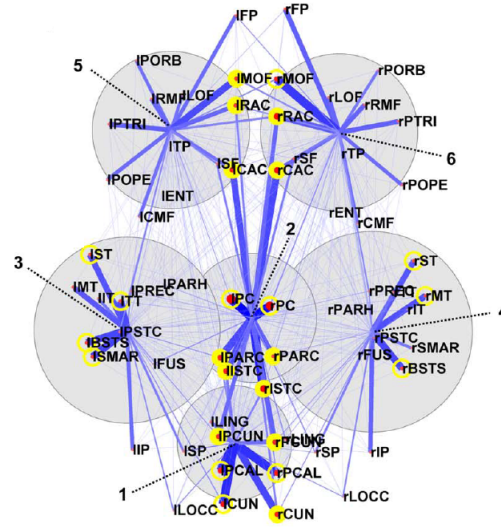
#### 4.2.2 Modularità e hubs

Dagli studi precedenti, emerge che l'architettura di piccolo mondo pervade le reti cerebrali strutturali. Una diretta conseguenza di ciò, come anche mostrato quantitativamente in [39], è la presenza di moduli, ossia di clusters di nodi (regioni) densamente interconnessi e legati tra loro da hubs: un altro segno particolare delle reti cerebrali è quindi la modularità, che a causa della natura multiscala della connettività strutturale, tende ad esprimersi in maniera gerarchica. Il carattere della modularità si configura essere peculiarità del connettoma ancora da mappare per diverse ragioni: alcune di carattere più generale valide per tutte le reti che processano informazioni, in cui sono ovviamente comprese quelle cerebrali; altre invece di carattere prettamente neuroscientifico [26]. In primis, un sistema modulare risulta più robusto agli attacchi e alle situazioni di vulnerabilità (cfr. sezione 2.3.3). Inoltre, in questo tipo di architettura, l'alto clustering promuove la formazione di aree

specializzate dal punto di vista funzionale, con un netto risparmio “energetico” nella lunghezza dei cammini, mentre il cammino caratteristico relativamente basso rende comunque possibile l’esecuzione di processi integrativi, che coinvolgono diverse funzioni della rete. Ancora, da un punto di vista “evolutivo”, l’architettura modulare permette alla rete di adattarsi ad un ambiente che richiede alla rete lo svolgimento di attività in continuo cambiamento. Infatti, l’esistenza di moduli destinati a svolgere precise funzioni di base permette al network di poter soddisfare richieste articolate sempre diverse, proprio grazie alla combinazione dinamica delle funzionalità basilari dei singoli moduli. In aggiunta a tali motivazioni di carattere generale, esistono importanti evidenze di tipo biologico della localizzazione anatomica di alcune funzioni specifiche nel cervello umano, supportate anche da ragioni di carattere genetico.

Nell’ambito dei 4 principali studi sulla connettività strutturale umana, se ne distinguono due nei quali è stato affrontato il problema della modularità. Partendo dal network ottenuto in [35] analizzando la correlazione tra gli spessori corticali di diverse aree della corteccia, Chen et al.(2008) [42] hanno applicato l’algoritmo di Girvan e Newman, il metodo GN (Eq. 1.12), per rilevare la natura modulare del network ricavato. Tramite l’algoritmo sono stati identificati 6 moduli, composti da un numero di nodi variabile tra 4 e 10: ciascun raggruppamento di nodi risulta corrispondere a gruppi di aree corticali che si distinguono per la funzione svolta come ad esempio visione, linguaggio e movimento.

Analogamente, lo studio di Hagmann [39], sempre con l’uso del metodo GN, rivela la presenza di 6 moduli (Fig. 19), configurazione rispetto alla quale il valore della modularità  $Q$  ha mostrato un picco. La figura mostra una vista dorsale: i sei moduli sono mostrati come cerchi estesi a seconda del numero di nodi che li compongono. In Fig. 19 compaiono anche gli hubs, distinti tra quelli connettori indicati con un pallino giallo pieno, per i quali l’indice di partecipazione  $P$  (Eq. 1.15) è maggiore di 0.3 e quelli provinciali indicati con un pallino giallo vuoto, con un indice  $P$  minore di 0.3. Quattro moduli sono localizzati ai lati in modo quasi simmetrico, nelle aree frontali e temporoparietali di ciascun emisfero, mentre gli altri due sono collocate lungo l’asse mediale, in posizione posteriore, centrati sui nodi che rappresentano il precuneo e la corteccia cingolata posteriore.



**Figura 19:** Hubs e modularità nello studio di Hagmann et al. 2008 [39]

L'algoritmo GN è stato anche applicato a porzioni ristrette del networks, all'interno di singoli moduli precedentemente individuati: l'analisi ha rivelato l'esistenza di sottomoduli annidati in maniera gerarchica nei precedenti, per i quali è sempre rintracciabile una coerenza funzionale. Un tale schema gerarchico è ritenuto essere costitutivo del cervello umano e di ogni rete neuroanatomica reale [3].

In generale quindi tutti gli studi concordano sul fatto che i moduli rappresentino gruppi di regioni funzionalmente correlate, in quanto, a causa dell'alto clustering, l'elaborazione dell'informazione avviene primariamente all'interno del modulo, preservando così la specializzazione di ciascuna comunità. Inoltre, emerge dall'analisi della densità delle connessioni che le regioni che costituiscono i moduli oltre ad essere topologicamente vicine sono anche spazialmente contigue. Una probabile ragione di ciò potrebbe essere che una tale tipologia struttura modulare permetta di incapsulare sistemi topologicamente complessi in uno spazio fisico relativamente piccolo. Inoltre, questa disposizione sembra essere vantaggiosa per minimizzare le lunghezze delle fibre intramodulari, ma implica anche che le connessioni intermodulari debbono essere lunghe e dispendiose dal punto di vista energetico [26].

Alla presenza dei moduli, è automaticamente associata la presenza di nodi hubs,

chiamati a gestire il flusso di informazioni tra i moduli: essi rappresentano aree altamente connesse e topologicamente centrali nel cervello. L'esistenza degli hubs è essenziale per mantenere l'integrazione globale di una rete costituita da comunità: per questa ragione, questo tipo di nodi sono punti di vulnerabilità, il cui malfunzionamento può generare problemi in tutta la rete, proprio a causa dei loro elevati centralità e grado. È necessario puntualizzare però che l'importanza di un hub all'interno di una rete cerebrale è data esclusivamente dal suo grado e dalla sua centralità: nemmeno in questo contesto essi possiedono uno speciale potere decisionale o autonomia di controllo su altri nodi [3].

Come è stato evidenziato in precedenza in Fig. 16, lo studio di Gong et al. [37] ha individuato 9 hubs principali: quelli maggiormente centrali tra questi secondo la grandezza della betweenness si sono rivelati essere il precuneo e il giro frontale superiore.

In maniera indipendente, anche lo studio di Hagmann [39] identifica numerosi hubs, una volta determinata la partizione ottimale in 6 moduli (Fig. 19). Tramite il criterio del coefficiente di partecipazione  $P$ , gli hubs vengono distinti in connettori e provinciali. Senza eccezioni, gli hubs connettori sono localizzati lungo l'asse mediale della corteccia, in particolar modo nella parte posteriore della corteccia che comprende il precuneo e la corteccia cingolata posteriore: una simile formazione di hubs forma un core strutturale. Gli hubs provinciali invece popolano maggiormente i 4 moduli laterali.

L'attenzione verso la ROI che rappresenta il precuneo è un successo della scienza delle reti applicata alla connettività strutturale del cervello: prima del suo utilizzo, il coinvolgimento di questa regione in un vasto numero di processi cognitivi come l'immaginazione, la memoria e lo stato di coscienza era piuttosto inspiegabile. L'utilizzo della teoria dei networks ha permesso di scoprire l'alta centralità topologica di questa area e l'elevato numero di cammini neuronali che la attraversano, fornendo una prima possibile interpretazione alla partecipazione del precuneo nei processi più vari. Inoltre, l'identificazione del precuneo come hub spiega come mai a seguito di sue lesioni, i danni cerebrali risultino piuttosto severi, rispetto a quelli riportati per il danneggiamento di altre aree [43].

## Conclusioni

In questo lavoro di tesi sono state passate in rassegna le modalità con le quali è possibile analizzare la connettività strutturale del cervello umano con l’approccio delle reti complesse. I risultati riportati sono solo alcune delle tessere di un puzzle chiamato connettoma umano che al momento è ben lontano da una completa “mappatura”. Comunque lo stato dell’arte degli studi consente già di delinearne le caratteristiche principali: il connettoma umano si configura essere una rete di piccolo mondo modulare e gerarchica e si sviluppa su più scale dimensionali. La mappatura completa delle connessioni cerebrali strutturali è quindi uno degli obiettivi principali della ricerca nelle neuroscienze perseguiti dai progetti quali lo *Human connectome project* promosso dagli Usa e lo *Human brain project* promosso dall’UE. Il raggiungimento di questo target pone diverse sfide alle discipline scientifiche coinvolte nel progetto e innanzitutto richiede che la fisica teorica fornisca modelli interpretativi dei fenomeni registrati sempre più pertinenti. Come mostrato in questo stesso lavoro, la versatilità e la concisione nel descrivere i più disparati sistemi complessi rappresentano i punti di forza della teoria delle reti: queste costituiscono quindi un modello elegante ed efficace per analizzare le caratteristiche integrative dei sistemi composti da molte parti interagenti, che invece quasi mai emergono da un’analisi riduzionistica. Di conseguenza, l’area della meccanica statistica che studia le reti complesse ha una specifica ragione d’essere proprio grazie alla versatilità ed all’estrema ampiezza delle possibilità di applicazioni della teoria. Infatti eventuali sviluppi modellistici non si applicherebbero soltanto allo studio del connettoma. Un ruolo altrettanto fondamentale nella costruzione del connettoma è assunto dallo sviluppo delle tecniche di imaging e di storage dei dati, che rappresentano attualmente un limite tecnologico ineludibile. La mappatura completa a livello cellulare sarà possibile solo mediante l’utilizzo di tecniche di imaging in vivo applicabili su vaste aree, caratterizzate da elevata sensibilità e risoluzione spaziale, e dalla possibilità della gestione dell’enorme quantità di dati raccolti. Nonostante le difficoltà di questo tipo di studio, innumerevoli sono i suoi vantaggi: così come alla decifrazione del genoma umano è conseguita una rivoluzione scientifica di portata epocale per l’impulso dato allo sviluppo delle

conoscenze biologiche e per la loro ricaduta nella pratica medica, allo stesso modo la conoscenza del connettoma potrà fornire una completa comprensione della funzionalità del cervello umano, legata strettamente alla sua struttura anatomica. Dalla conoscenza dettagliata degli aspetti “cinematici”, sarà possibile quindi risalire alla dinamica e ai modi evolutivi del sistema nervoso, come già accade adesso, su scale limitate. Le implicazioni di questo genere di approccio sono evidenti: un modello completo della funzionalità cerebrale come “rete” permetterebbe di scoprire eventuali anomalie di connettività, che da “semplici” discrepanze rispetto al modello fisico, si rivelerebbero come segno di una patologia. Inoltre un modello renderebbe possibile la predizione delle reazioni del sistema nervoso nel suo complesso in relazione a determinati stimoli, delle sue vulnerabilità in caso di pericolo, nonché delle capacità plastiche di auto adattamento a seguito a danneggiamento. Oltre alla funzione diagnostica, la costruzione del connettoma potrà fornire nuove chiavi interpretative per la comprensione del cervello in termini evoluzionistici. I modelli delle reti infatti permettono di ricostruire l'evoluzione di un sistema a partire dall'architettura che esso assume: da questo punto di vista, si potrebbe dare un fondamento teorico alle proprietà dei circuiti neuronali in termini di economia energetica nella formazione delle connessioni e nella necessità di contenere una rete complessa in uno spazio fisico comunque limitato. In ultima analisi, gli strumenti interpretativi offerti dalla teoria delle reti complesse a partire già dai risultati di connettività strutturale della corteccia presentati in questo lavoro sono numerosi, estremamente promettenti e si configurano come possibile framework teorico per la ricerca neuroscientifica.

## Riferimenti bibliografici

- [1] Stefano Boccaletti, Vito Latora, Yamir Moreno, Martin Chavez, and D-U Hwang. Complex networks: Structure and dynamics. *Physics reports*, 424(4):175–308, 2006.
- [2] Euler. Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum imperialis petropolitanae*, 1736.
- [3] Olaf Sporns. *Networks of the Brain*. MIT press, 2011.
- [4] Alain Barrat, Marc Barthélemy, Romualdo Pastor-Satorras, and Alessandro Vespignani. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752, 2004.
- [5] Marc Barthélemy, Alain Barrat, Romualdo Pastor-Satorras, and Alessandro Vespignani. Characterization and modeling of weighted networks. *Physica a: Statistical mechanics and its applications*, 346(1):34–43, 2005.
- [6] M Boguñá, R Pastor-Satorras, and A Vespignani. Statistical mechanics of complex networks lecture notes in physics volume 625. 2003.
- [7] Romualdo Pastor-Satorras, Alexei Vázquez, and Alessandro Vespignani. Dynamical and correlation properties of the internet. *Physical review letters*, 87(25):258701, 2001.
- [8] Mark EJ Newman. Assortative mixing in networks. *Physical review letters*, 89(20):208701, 2002.
- [9] Duncan J Watts and Steven H Strogatz. Collective dynamics of small-world networks. *nature*, 393(6684):440–442, 1998.
- [10] Mark EJ Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.

- [11] Shai S Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature genetics*, 31(1):64–68, 2002.
- [12] Ron Milo, Shalev Itzkovitz, Nadav Kashtan, Reuven Levitt, Shai Shen-Orr, Inbal Ayzenshtat, Michal Sheffer, and Uri Alon. Superfamilies of evolved and designed networks. *Science*, 303(5663):1538–1542, 2004.
- [13] Michelle Girvan and Mark EJ Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, 2002.
- [14] Mark EJ Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113, 2004.
- [15] Vito Latora and Massimo Marchiori. Economic small-world behavior in weighted networks. *The European Physical Journal B-Condensed Matter and Complex Systems*, 32(2):249–263, 2003.
- [16] Roger Guimera, Marta Sales-Pardo, and Luis AN Amaral. Classes of complex networks defined by role-to-role connectivity profiles. *Nature physics*, 3(1):63–69, 2007.
- [17] Stanley Wasserman. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.
- [18] D.B. West. *Introduction to graph theory*. Prentice Hall, 2001.
- [19] Paul Erdos and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5:17–61, 1960.
- [20] B. Bollobás. Random graphs. In *Modern Graph Theory*, volume 184 of *Graduate Texts in Mathematics*, pages 215–252. Springer New York, 1998.

- [21] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64:026118, Jul 2001.
- [22] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.
- [23] Mark D Humphries and Kevin Gurney. Network small-world-ness: a quantitative method for determining canonical network equivalence. *PloS one*, 3(4):e0002051, 2008.
- [24] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.
- [25] Luis A Nunes Amaral, Antonio Scala, Marc Barthelemy, and H Eugene Stanley. Classes of small-world networks. *Proceedings of the National Academy of Sciences*, 97(21):11149–11152, 2000.
- [26] David Meunier, Renaud Lambiotte, and Edward T Bullmore. Modular and hierarchically modular organization of brain networks. *Frontiers in neuroscience*, 4, 2010.
- [27] Herbert A Simon. Near decomposability and complexity: How a mind resides in a brain. In *SANTA FE INSTITUTE STUDIES IN THE SCIENCES OF COMPLEXITY-PROCEEDINGS VOLUME-*, volume 22, pages 25–25. ADDISON-WESLEY PUBLISHING CO, 1995.
- [28] Ricard V Solé and Sergi Valverde. Information theory of complex networks: On evolution and architectural constraints. In *Complex networks*, pages 189–207. Springer, 2004.
- [29] Olaf Sporns. Structure and function of complex brain networks. *Dialogues in clinical neuroscience*, 15(3):247, 2013.

- [30] Olaf Sporns, Giulio Tononi, and Rolf Kötter. The human connectome: a structural description of the human brain. *PLoS computational biology*, 1(4):e42, 2005.
- [31] Olaf Sporns. The human connectome: a complex network. *Annals of the New York Academy of Sciences*, 1224(1):109–125, 2011.
- [32] John G White, Eileen Southgate, J Nichol Thomson, and Sydney Brenner. The structure of the nervous system of the nematode *caenorhabditis elegans*. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 314(1165):1–340, 1986.
- [33] Winfried Denk and Heinz Horstmann. Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS biology*, 2(11):e329, 2004.
- [34] Jean Livet, Tamily A Weissman, Hyuno Kang, Ju Lu, Robyn A Bennis, Joshua R Sanes, and Jeff W Lichtman. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature*, 450(7166):56–62, 2007.
- [35] Yong He, Zhang J Chen, and Alan C Evans. Small-world anatomical networks in the human brain revealed by cortical thickness from mri. *Cerebral cortex*, 17(10):2407–2419, 2007.
- [36] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198, 2009.
- [37] Gaolang Gong, Yong He, Luis Concha, Catherine Lebel, Donald W Gross, Alan C Evans, and Christian Beaulieu. Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cerebral cortex*, 19(3):524–536, 2009.

- [38] Patric Hagmann, Maciej Kurant, Xavier Gigandet, Patrick Thiran, Van J Wedeen, Reto Meuli, and Jean-Philippe Thiran. Mapping human whole-brain structural networks with diffusion mri. *PloS one*, 2(7):e597, 2007.
- [39] Patric Hagmann, Leila Cammoun, Xavier Gigandet, Reto Meuli, Christopher J Honey, Van J Wedeen, and Olaf Sporns. Mapping the structural core of human cerebral cortex. *PLoS biology*, 6(7):e159, 2008.
- [40] Olaf Sporns, Giulio Tononi, and Gerald M Edelman. Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cerebral Cortex*, 10(2):127–141, 2000.
- [41] Danielle Smith Bassett and ED Bullmore. Small-world brain networks. *The neuroscientist*, 12(6):512–523, 2006.
- [42] Zhang J Chen, Yong He, Pedro Rosa-Neto, Jurgen Germann, and Alan C Evans. Revealing modular architecture of human brain structural networks by using cortical thickness from mri. *Cerebral cortex*, 18(10):2374–2381, 2008.
- [43] Antonio Damasio. The feeling of what happens. 1999. *Harcourt Brace, New York*.

## Ringraziamenti

*"Ma le occasioni della vita stupiscono mai abbastanza nella loro frammentarietà che poi un bel giorno miracolosamente si salda in una sottile e delicata vibrazione che riaccorda e riannoda e uniforma il tono di diversi percorsi e allora, nonostante i dolori e le precarietà dei nostri anni giovanili la vita sembra rivelarsi come una misteriosa e armonica frequenza che schiude il senso e fa capire; e allora in quell'attimo abbagliante tutto pare ricomporsi nella gioia di sentirsi finalmente presenti agli occhi della propria storia, la pazzesca consapevolezza di trarre a sé tutti i fili intrigati e sparsi del proprio passato come appunto sta succedendo a me, ora, nella luce calda di questa città in cui ogni giorno, miracolosamente, incontro qualche personaggio di questa storia che vi sto raccontando."* (Pier Vittorio Tondelli, Pao Pao)

Al compimento di un percorso "complesso e nodoso" come quello universitario, è necessario ringraziare quelle figure che a vario titolo hanno percorso con me un pezzo di strada. In primo luogo, vorrei ringraziare il prof. Sebino Stramaglia per la professionalità, la pazienza ed attenzione che mi ha dedicato nonostante i problemi logistici e le mie ansie pressanti.

In secondo luogo, ma non per importanza, vorrei esprimere un sentito "grazie" ai miei familiari: ai miei genitori, per il loro sostegno continuo e fiducia incondizionata nonostante i miei spigoli e discontinuità, e a Francesco, che considero un fratello.

Desidero ringraziare gli amici, tutti, i quali hanno avuto un ruolo non marginale in questi anni: la mia gratitudine a loro va per avermi ascoltato, placato le mie ansie e fatto distrarre nei momenti peggiori. In particolare, vorrei ringraziare coloro che da colleghi sono diventati amici: in primo luogo Sissi, compagna di esami, gelati e malefatte. A lei una menzione speciale per avermi insegnato a fare le fotocopie, per avermi messo in crisi con i suoi dubbi, per avermi ricordato scadenze che altrimenti sarebbero state dimenticate e per essersi rivelata inoltre una vera amica. Un'altra menzione a Maurangelo e Gianfranco, per aver condiviso con me risate interminabili, questioni scientifiche, dissertazioni filosofiche, birre e concerti. Un'ultima menzione a Gaetano, che ha segnato gli ultimi mesi del mio percorso mostrando un'incredibile pazienza alle mie incessanti e spesso esagerate preoccupazioni.

Vorrei esprimere la mia gratitudine inoltre ai miei amici storici, per l'affetto e calore che mi hanno dimostrato in tutte le diverse fasi della nostra amicizia, della mia e della loro esistenza condivisa: un grazie a Roberto e Giovanna, Daniele, Piergiovanni e Ciccio. Un ringraziamento a Vincenzo S., per la complicità e la stima dimostratami negli anni del Conservatorio ed oltre.

Un ringraziamento, in ultimo, sento di farlo verso gli amici con i quali condivido serate, risate, psicodrammi, litigi per aver tollerato i miei no e le mie assenze: sono parte dei miei “incontri miracolosi”, forse uno tra i regali più grandi ed inaspettati da me ricevuti ed a loro sento di esprimere grande riconoscenza. Non li cito uno ad uno ma ciascuno di loro si riconoscerà: una nota speciale però per Pierpaolo, per l'amicizia incondizionata e da me immeritata e per l'elefantiaca pazienza nell'avermi aiutato con L<sup>A</sup>T<sub>E</sub>X.