

PROYECTO INTEGRADOR DE LA CARRERA DE INGENIERÍA EN TELECOMUNICACIONES

MEDIDAS DE COMUNICACIÓN EN REDES CEREBRALES Y SU POSIBLE ROL EN LA CLASIFICACIÓN DE PATOLOGÍAS

Tomás Gerardo Hüttebräucker
Estudiante

Dr. Sebastián Risau Gusmán
Director

Miembros del Jurado
Damían Hernandez Lahme
Marcelo Kuperman

4 de Noviembre de 2021

Departamento de Física Médica

Instituto Balseiro
Universidad Nacional de Cuyo
Comisión Nacional de Energía Atómica
Argentina

Índice de símbolos

Índice de contenidos

Índice de símbolos	v
Índice de contenidos	vii
Resumen	ix
Abstract	xi
1. Introducción	1
1.1. Cerebro humano	1
1.1.1. Neuronas	1
1.1.2. Anatomía del cerebro humano	1
1.1.3. Imágenes cerebrales por resonancia magnética nuclear	2
1.2. Conectomas	3
1.2.1. Regiones de Interés	3
1.2.2. Tractografía	4
1.2.3. Construyendo un conectoma	4
1.3. Análisis de redes cerebrales	4
1.4. Objetivos y materiales	4
1.4.1. Materiales y herramientas	6
2. Redes	7
2.1. Clasificación de grafos y matriz de adyacencia.	7
2.2. Comunicación en redes	8
2.2.1. Ruteo	8
2.2.2. Difusión	10
2.2.3. Cálculo de caminos mas cortos	10
2.3. Medidas de comunicación en redes	12
2.3.1. Eficiencia de Ruteo	12
2.3.2. Eficiencia de Difusión	14
2.3.3. Comunicabilidad	14
2.3.4. Distancia de K-caminos	15
2.3.5. Centralidad de nodos y aristas	16
2.3.6. <i>Search information</i>	16
2.3.7. Matriz de acoplamiento	16
A. Apéndice I	17
Bibliografía	19

Publicaciones asociadas	21
Agradecimientos	23

Resumen

Este es el resumen en castellano.

La tesis debe reflejar el trabajo desarrollado, mostrando la metodología utilizada, los resultados obtenidos y las conclusiones que pueden inferirse de dichos resultados.

Palabras clave: FORMATO DE TESIS, LINEAMIENTOS DE ESCRITURA, INSTITUTO BALSEIRO

Abstract

This is the title in English:

The thesis must reflect the work of the student, including the chosen methodology, the results and the conclusions that those results allow us to draw.

Keywords: THESIS FORMAT, TEMPLATES, INSTITUTO BALSEIRO

Capítulo 1

Introducción

Se suele decir que el cerebro es la red más compleja que conoce el humano. Un cerebro humano está compuesto por 100 mil millones (10^{11}) de neuronas conectadas por 100 billones (10^{14}) de sinapsis, que están organizadas en múltiples escalas espaciales y son funcionalmente interactivas en múltiples escalas temporales. Las neuronas por si mismas no pueden llevar a cabo las complejas funciones de un cerebro, pero es cuando estas se conectan y organizan un sistema nervioso cuando los pensamientos, la memoria, la conciencia y los sentimientos se vuelven realidad. Muchas patologías neuronales están asociadas a daños en la red cerebral. La ubicación, extensión y caracterización de esta alteración puede permitir predecir la gravedad de la patología así como mejorar las probabilidades de recuperación del paciente. Teniendo en cuenta esto, no es sorpresa que entender la conectividad de la red cerebral y como surgen las diversas funciones cerebrales a partir de esta sea uno de los principales objetivos de la neurociencia.

Las técnicas de mapeo cerebral modernas (dMRI, fMRI, EEG, MEG), permiten obtener un mapa de las conexiones anatómicas o funcionales del cerebro. Para caracterizar estos mapeos, el análisis de redes complejas es el enfoque más utilizado. En este capítulo se presentará primero una breve explicación de la anatomía cerebral, luego se introducirá el concepto de conectoma (mapeo de conexiones cerebrales) y como se obtienen conectomas estructurales usando imágenes por tensor de difusión. Se dará también una breve explicación sobre redes y algunos resultados del análisis de las mismas en el campo de la neurociencia y por último se resumirán los objetivos del proyecto integrador.

1.1. Cerebro humano

1.1.1. Neuronas

Las neuronas son células capaces de recibir, procesar y transmitir información a través de señales químicas y eléctricas mediante conexiones llamadas sinapsis. Son el principal componente del sistema nervioso en la mayoría de los animales y particularmente en los humanos. Anatómicamente, están formadas por el soma o cuerpo celular, las dendritas y el axón. El axón es la parte de la neurona encargada de transmitir la información hacia otras células, mientras que es en el soma donde la señal se genera y propaga. En la figura 1.1 se muestra un esquema de la estructura anatómica de una neurona.

1.1.2. Anatomía del cerebro humano

El cerebro humano está formado por una capa exterior llamada corteza cerebral, compuesta principalmente por materia gris, y por un centro formado por materia blanca. La materia gris está formada

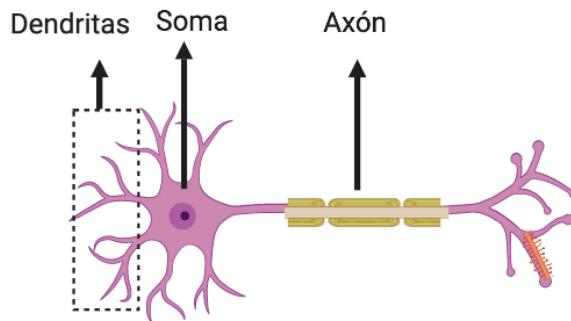


Figura 1.1: Estructura anatómica de una neurona

mayormente por los soma, dendritas y sinapsis de las neuronas, mientras que la materia blanca esta formada principalmente por axones. Conociendo las funciones de los axones y el soma, dendritas y sinapsis, se puede pensar, de forma simplificada, al cerebro como conjuntos de neuronas (la materia gris de la corteza) comunicadas entre si por un tramado de fibras nerviosas (la materia blanca del núcleo). En la figura 1.2 se muestra una representación de la estructura cerebral humana formada por materia gris y materia blanca.

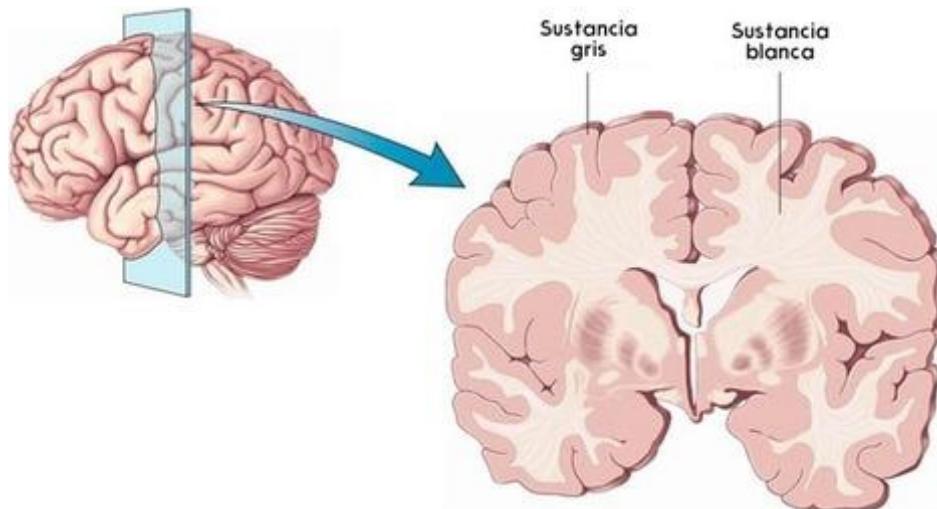


Figura 1.2: Representación de la estructura cerebral humana

1.1.3. Imágenes cerebrales por resonancia magnética nuclear

La resonancia magnética nuclear, RMN o simplemente RM, es una herramienta que permite obtener imágenes de la anatomía y procesos fisiológicos del cuerpo humano de forma no invasiva y sin hacer uso de radiación ionizante. Las imágenes se obtienen al aplicar campos magnéticos fuertes, gradientes de campos magnéticos y ondas de radio y observar como interactúan con los diferentes tejidos del cuerpo. Entre los tipos de imágenes por RM que se pueden obtener del cerebro hay tres que serán relevantes para el desarrollo de este trabajo, las imágenes por tensor de difusión, las imágenes por RM funcional y las imágenes por RM volumétricas.

Imágenes por tensor de difusión

Las imágenes por tensor de difusión, ITD, permiten medir el tensor de difusión del agua en diferentes partes del cerebro, correspondientes a un píxel o voxel en la imagen. Debido a que el agua

tiende a difundir en la dirección de los tractos axonales del cerebro, conocer el tensor de difusión del agua en una zona dada, brinda información sobre la forma en que están orientados los axones en esa zona. Ya que normalmente la resolución de la imagen no es del orden del tamaño de los axones, el tensor de difusión medido en un lugar del espacio es un promedio de la contribución de varias fibras axonales presentes en ese lugar. Las ITD permiten ver como están conectadas estructuralmente las diferentes zonas del cerebro.

Imágenes por RM funcional

Las imágenes por RM funcional, IRMf, permiten medir cambios en la actividad neuronal en las diferentes partes del cerebro al detectar cambios asociados al flujo sanguíneo. Se puede medir la actividad neuronal como respuesta a un estímulo o también la actividad neuronal base de cada sujeto (fMRI en estado de reposo). Este tipo de imágenes sirve para entender las funciones que cumplen diferentes regiones del cerebro y como interactúan entre ellas en diferentes situaciones.

Imágenes por RM volumétricas

Las imágenes por RM volumétricas permiten capturar la estructura tridimensional del interior del cuerpo al juntar múltiples imágenes bidimensionales. Este tipo de imágenes es vital para poder identificar las diferentes regiones del cerebro, lo que permite un análisis mas profundo utilizando las IRMf o las ITD explicadas anteriormente.

1.2. Conectomas

Un conectoma es una representación de las conexiones que existen entre diferentes regiones del cerebro humano. Las conexiones pueden ser:

- Anatómicas o estructurales: representan las conexiones físicas (de materia blanca) entre regiones del cerebro, se obtienen a partir de imágenes por tensor de difusión.
- Funcionales: describen dependencias estadísticas entre las actividades neuronales entre regiones del cerebro, se obtienen a partir de las imágenes por resonancia magnética funcional.

Para obtener un conectoma estructural es necesario dividir el cerebro en diferentes regiones de interés y obtener una medida de las fibras axonales. Las regiones de interés del cerebro se definen haciendo ajustando las imágenes por RM volumétricas de acuerdo a una división espacial estándar del cerebro llamada atlas. Las fibras axonales se pueden caracterizar al aplicar un algoritmo de reconstrucción de tractos a una ITD.

1.2.1. Regiones de Interés

Definir la forma en la que se va a dividir el cerebro en regiones de interés discretas y espacialmente contiguas no es una tarea simple. En general, se puede dividir el cerebro en diferentes partes de acuerdo a la ubicación, anatomía y funciones con las que están relacionadas. La división obtenida depende también de la escala que se utilice, en una escala mas pequeña se pueden subdividir las regiones de interés presentes en una escala mas grande. Para obtener la división en una imagen dada se usan los atlas cerebrales, que permiten mapear múltiples voxels de la imagen del cerebro a diferentes regiones de interés de acuerdo a una división modelo basándose en la posición, forma, tamaño y estructura anatómica obtenidas con una imagen por RM volumétrica.

1.2.2. Tractografía

La tractografía es la reconstrucción de tractos de fibras axonales a partir de las mediciones del tensor de difusión del agua. Hay tres pasos importantes a seguir para realizar una tractografía:

- Primero hay que definir la estimación de la orientación axonal en un píxel o voxel, para esto se puede usar el del tensor de difusión completo o simplemente la dirección principal del tensor de difusión. El último enfoque es mas directo pero inevitablemente se pierde información, sobre todo sobre la fiabilidad de esta dirección principal. Cuanto mas anisotrópico sea el tensor de difusión, mas confiable es la dirección principal del mismo como estimación de la orientación axonal. En el caso extremo en el que el tensor de difusión es completamente isotrópico (misma difusión en todas las direcciones), no se puede confiar en la dirección obtenida, ya que sería simplemente producto del ruido. En la figura 1.3a se muestra el contraste entre las dos opciones para estimar la orientación axonal.
- El segundo paso es la propagación de los tractos basados en el campo de orientación axonal elegido, para esto hay varios métodos basados en interpolación o simulaciones.
- El tercer y último paso consiste en definir el final de un tracto, para esto los dos criterios mas usados son el de baja anisotropía y el de curva cerrada. En el criterio de baja anisotropía se asume que no hay una población coherente de fibras y en el de curva cerrada que la resolución de la imagen no es suficiente para detectar una curva de ese tipo. En la figura 1.3b se muestra la forma en que termina la propagación de tractos de acuerdo a los dos criterios mencionados.

1.2.3. Construyendo un conectoma

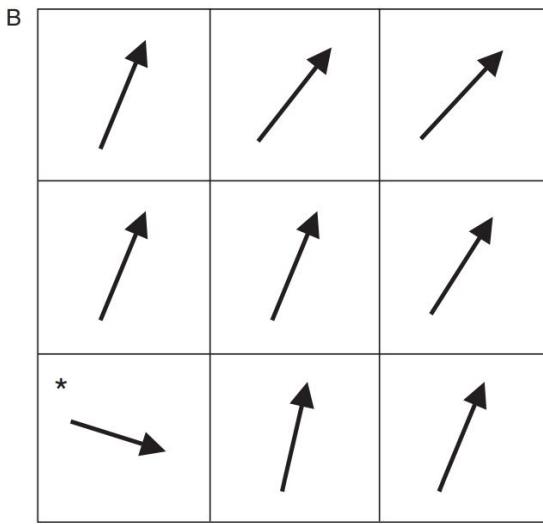
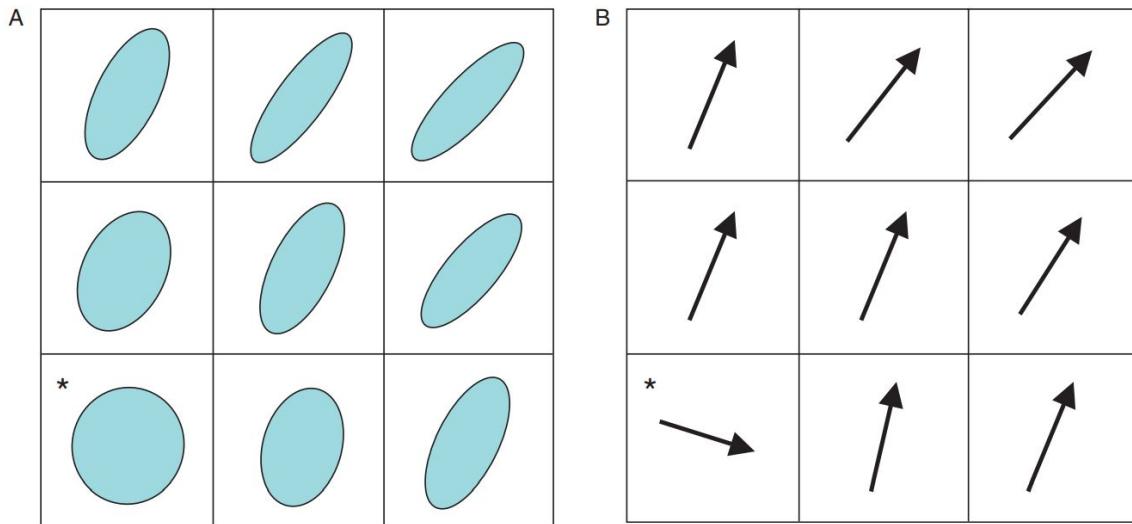
Una vez hecha la tractografía y la parcelación queda reconstruir las conexiones entre las diferentes regiones de interés. En esta instancia es común el uso de algoritmos de posprocesado de tractografía que permiten obtener un número de fibras reconstruidas que sea valido como marcador biológico como el algoritmo SIFT2 [2].

1.3. Análisis de redes cerebrales

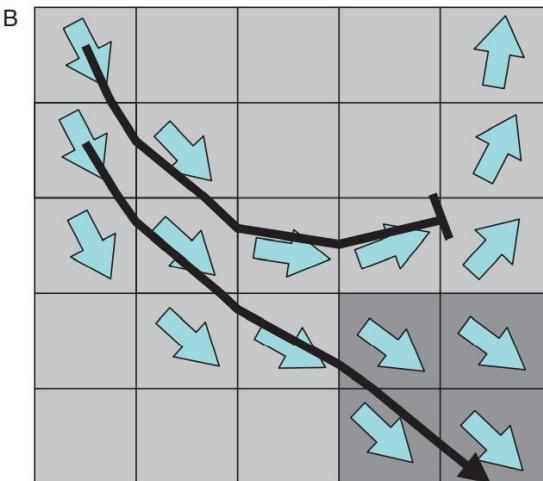
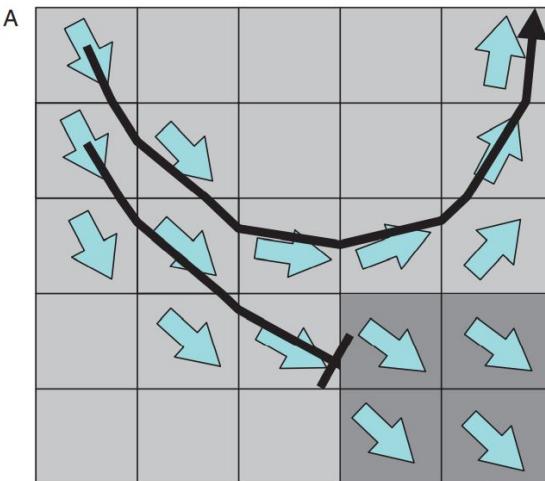
Las redes o grafos son una forma de modelar matemáticamente diversos sistemas complejos formados por elementos fundamentales conectados entre si. Una red esta formada por un conjunto de nodos y conexiones o aristas entre pares de nodos. En el caso de redes cerebrales, estas son representadas por los conectomas, los nodos representan a las regiones de interés y las aristas a las conexiones estructurales o funcionales que fueron obtenidas entre esas regiones con los métodos explicados anteriormente. El análisis de redes cerebrales ha llevado a diversos resultados importantes, Goñi et al. [3] pudieron predecir con bastante precisión la conectividad funcional utilizando medidas de comunicación sobre conectomas estructurales y Crofts et al. [4] encontraron diferencias sustanciales en las redes estructurales de pacientes que sufrieron un accidente cerebrovascular . Estos resultados demuestran que el análisis de redes cerebrales puede ser una herramienta de suma importancia para caracterizar patologías y entender mejor como funciona el cerebro.

1.4. Objetivos y materiales

El primer objetivo del proyecto es caracterizar la validez o confiabilidad de un conectoma estructural obtenido. Debido a las diferentes fuentes de ruido y los diferentes métodos usados para obtener un



(a) Representación de los campos de estimación de orientación axonal tensorial (A) y vectorial (B). El campo vectorial está compuesto por las direcciones principales del campo tensorial en cada píxel o voxel. En el caso en el que el tensor sea completamente isotrópico (píxel marcado con asterisco), el vector asociado no es confiable.



(b) Reconstrucción de tractos a partir del campo de orientación axonal obtenido usando los criterios de baja anisotropía (A) y de curva cerrada (B). Los píxeles más oscuros se corresponden con una anisotropía baja.

Figura 1.3: Representación de los diferentes criterios posibles a la hora de realizar una tractografía.[1]

conectoma, muchas veces se consideran a las conexiones débiles obtenidas como conexiones espurias y se eliminan de la red previo al análisis de la misma. No esta establecido que esta práctica, conocida como podado o *pruning*, sea correcta y aporte algún valor apreciable al análisis de redes cerebrales, es por esto que evaluar el impacto que tiene sobre diferentes propiedades topológicas de las mismas es de utilidad para buscar un consenso al respecto.

El segundo y principal objetivo de este trabajo es el de aplicar diferentes medidas de comunicación a redes cerebrales estructurales para detectar diferencias entre una población de pacientes con migraña crónica y una población de personas sanas. Las medidas a probar son medidas usadas en estudios previos que han permitido caracterizar o predecir propiedades biológicas o matemáticas de redes cerebrales. En este sentido también se busca explorar diferentes filosofías de comunicación comúnmente usadas en redes cerebrales, ruteo o difusión.

1.4.1. Materiales y herramientas

El material con el que se trabajó en el presente proyecto consiste en matrices de conexión estructural correspondientes a 72 pacientes con migraña episódica, 69 pacientes con migraña crónica y 62 sujetos de control. Todos los datos fueron recolectados por Planchuelo-Gómez et al [5]. Los pacientes fueron reclutados en la unidad de dolores de cabeza del Hospital Clínico Universitario de Valladolid (Valladolid, España) y los sujetos de control fueron reclutados entre colegas y avisos en el hospital y la universidad.

El análisis de las redes cerebrales se realizó en el lenguaje de programación **Python 3** utilizando las librerías **Networkx** y **Numpy** principalmente.

Capítulo 2

Redes

Las redes, también llamadas grafos, son la representación matemática de un sistema complejo que involucra elementos comunicados entre si. Un grafo está compuesto por un conjunto de nodos y un conjunto de aristas entre nodos. Los nodos representan los elementos fundamentales del sistema y las aristas representan las conexiones entre estos elementos. En el caso de una red basada en un conectoma estructural, los nodos representan las regiones de interés definidas y las aristas las conexiones de tractos axonales obtenidas mediante una tractografía.

En este capítulo se dará una introducción a teoría de grafos, primero viendo la clasificación y su representación mediante la matriz de adyacencia. Luego se explicarán el concepto de comunicación en redes, mostrando los dos modelos predominantes en el análisis de redes cerebrales y por último se introducirán las medidas utilizadas en este proyecto para caracterizar la comunicación en estas redes.

2.1. Clasificación de grafos y matriz de adyacencia.

Los grafos se clasifican de acuerdo al tipo de conexiones tengan. Las conexiones pueden ser binarias, es decir que existe o no existe conexión, o pueden tener un peso asociado, en este caso, si la conexión existe, el peso puede significar la fuerza o importancia de la conexión. También pueden diferenciarse las conexiones en caso de que sean no dirigidas, es decir que la conexión que hay entre dos nodos es simétrica, o dirigidas, en este caso las conexiones son asimétricas. En la figura 2.1 se muestran representaciones gráficas de los diferentes tipos de grafos de acuerdo a sus conexiones.

Una de las formas más simples y útiles de representar un grafo es la matriz de adyacencia. La matriz de adyacencia A define la topología de un grafo representando a los nodos como filas y columnas y a las aristas como los valores presentes en la matriz. Es decir, el valor a_{ij} (fila i y columna j de la matriz) representa la conexión desde el nodo i hacia el nodo j del grafo, si $a_{ij} = 1$ quiere decir que existe una arista que une al nodo i con el j , si $a_{ij} = 0$ no hay arista desde el nodo i hacia el nodo j . A dos nodos que estén conectados directamente (que exista una arista entre ellos) se los conoce como nodos vecinos. Es normal hacer una diferencia entre la matriz de adyacencia A y la matriz de pesos W , mientras que la matriz de adyacencia representa la existencia o no de aristas, el elemento w_{ij} de la matriz de pesos representa las aristas con sus pesos. En el caso de un grafo binario $w_{ij} = a_{ij}$ ya que el peso de una arista siempre es 1, mientras que en un grafo pesado no necesariamente se cumple esta igualdad. En el caso de un grafo no dirigido $a_{ij} = a_{ji}$ y $w_{ij} = w_{ji}$, por la simetría de la conexión, mientras que en un grafo dirigido esta igualdad no necesariamente se cumple. En la figura 2.1 se muestran los diferentes tipos de grafos con su representación gráfica y su matriz de pesos al lado.

Definidas las matrices de adyacencia y pesos, se puede definir el grado y el grado pesado de un nodo para grafos no dirigidos binarios o pesados :

$$k_i = \sum_{j \in N} a_{ij} \quad (2.1)$$

$$s_i = \sum_{j \in N} w_{ij} \quad (2.2)$$

donde N es el conjunto de todos los nodos del grafo, k_i es el grado del nodo i y s_i es el grado pesado o “fuerza” del nodo i .

2.2. Comunicación en redes

Como se dijo en el Capítulo 1, lo que hace posible que el cerebro pueda cumplir todas sus complejas funciones es la comunicación y organización que surge entre las neuronas que lo componen. Es por esto que es importante entender que información se puede obtener de un grafo sobre la comunicación del sistema que representa, en este caso la red cerebral. La capacidad de comunicarse de dos nodos de una red esta condicionada por los caminos que los conecten. Un camino es una secuencia de aristas ρ y nodos Ω que une a dos nodos de una red. Se dice que un camino es sin bucles cuando ninguno de los nodos que lo conforman esta repetido. Una característica importante de un camino es su longitud, la longitud de un camino es la suma de las distancias de las aristas que lo componen. En un grafo binario, la distancia asociada a una arista es 1 y la longitud de un camino es el número de aristas atravesadas. En un grafo pesado, la distancia asociada a una arista es alguna función del peso de la misma, la forma de elegir esta función depende de la red y no es trivial. En redes cerebrales estructurales, donde el peso de las conexiones representa el número de tractos axonales que unen a dos regiones, es razonable elegir una métrica decreciente con el peso de la conexión, es decir, que mientras mas grande sea el peso, menor sea la distancia asociada a esa arista. De esta forma, si dos regiones están unidas por una gran cantidad de tractos de materia blanca (un peso alto) la distancia entre esas regiones es baja (la función evaluada en ese peso es chica). Para medir la forma en que dos nodos de una red se comunican existen dos modelos diferentes: ruteo y difusión.

2.2.1. Ruteo

El modelo de ruteo supone que la propagación de una señal de entre dos nodos ocurre a través del camino mas corto que los une. El camino mas corto entre las regiones i y j de una red se denota como π_{ij} , mientras que la secuencia de aristas y nodos que lo conforman son ρ_{ij} y Ω_{ij} respectivamente. La longitud de este camino, l_{ij} , viene dada por:

$$l_{ij} = \sum_{w_{nm} \in \rho_{ij}} f(w_{nm}), \quad (2.3)$$

donde w_{nm} es el peso de la arista que une al nodo n con el nodo m. En el caso de una red binaria, f es la identidad y l_{ij} es el numero de aristas que componen al camino. La distancia entre dos nodos de la red se define como la longitud del camino mas corto que los une, de esta forma l_{ij} es la distancia entre el nodo i y el nodo j de la red. En la sección 2.2.3 se introduce el algoritmo de Dijkstra, usado para calcular el camino mas corto entre dos nodos de una red.

El uso de caminos mas cortos para propagar señales es claramente mas eficiente, al usar caminos cortos, la señal llega mas rápido y es menos probable que se pierda o que se llene de ruido. Sin embargo, esta suposición implica que cada elemento de la red, o neurona, tiene conocimiento sobre la topología

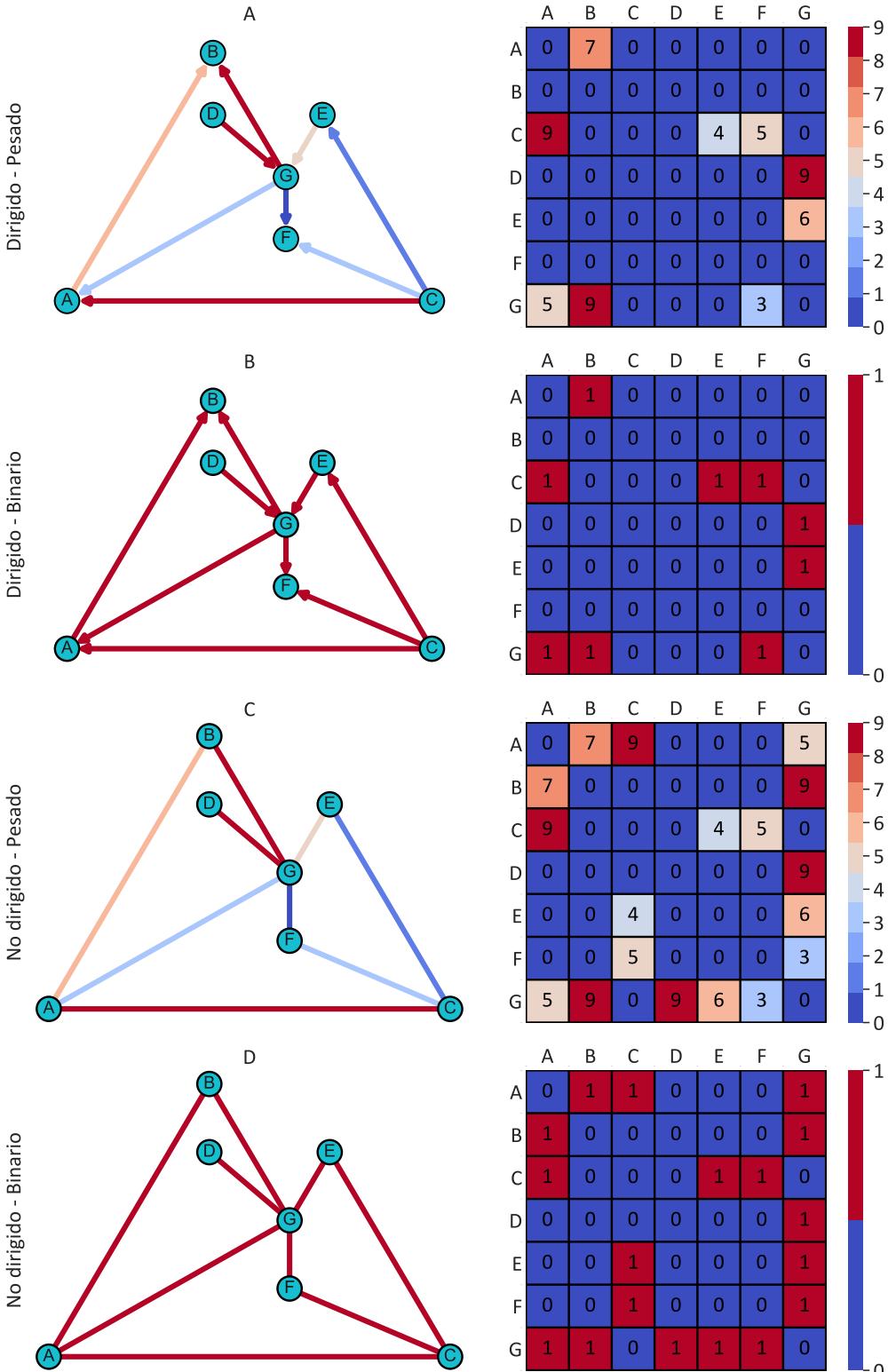


Figura 2.1: Representación gráfica de diferentes tipos de grafos. **(A)** Un grafo dirigido y pesado (izquierda), las flechas representan la direccionalidad de las conexiones y el color de las aristas representando el peso. En la matriz de pesos (derecha) se muestran los pesos de las conexiones con número y color. **(B)** El mismo grafo que en **(A)** pero binario (izquierda), las flechas siguen representando la direccionalidad y ahora el color es el mismo para todas las aristas. En la matriz de conectividad (derecha) las conexiones tienen valor 0 o 1. **(C)** Mismo grafo que en **(A)** pero no dirigido (izquierda), las conexiones ya no tienen direccionalidad, por eso no hay flechas. La matriz de pesos (derecha) es simétrica. **(D)** Versión binaria y no dirigida del grafo en **C** (izquierda), mismo color para todas las conexiones y ausencia de flechas. La matriz de pesos (derecha) es simétrica y con valores 0 o 1.

global de la misma. ¿Cómo puede saber una neurona cual es la ruta mas corta para enviar una señal a otra neurona? Es improbable que lo sepa y es por esto que surge otro modelo sobre transmisión de señales cerebrales, el modelo de difusión.

2.2.2. Difusión

En contraste con el modelo de ruteo, el modelo de difusión supone que las señales cerebrales se transmiten mediante un proceso de difusión. En un proceso de difusión las señales no viajan a través de los caminos mas cortos o a través de un camino en particular, si no que se propagan de forma dispersiva, a través de múltiples frentes, usando varios caminos diferentes a la vez y sin un destino único. La información enviada desde un nodo puede llegar a varios destinos a través de varios caminos diferentes, la difusión puede ser pensada como un proceso de *broadcasting*. Las señales enviadas siguen la lógica de un caminante aleatorio o *random walker* ya que no conocen la ruta a otro nodo y simplemente se transmiten de acuerdo a la topología local de la red.

En un contexto de propagación de señales por difusión, los caminos mas cortos no son los únicos relevantes para caracterizar que tan comunicados están dos nodos de la red. Todos los caminos que existan son sirven caracterizar la comunicación entre ellos. Sin embargo, la cantidad de caminos que existen entre dos nodos aumenta de forma factorial con el tamaño del grafo, por esto se suele usar el conjunto de los K caminos mas cortos. En la sección 2.2.3 se introducirá el algoritmo de Yen que sirve para obtener los K caminos mas cortos entre un nodo fuente y un nodo objetivo.

2.2.3. Cálculo de caminos mas cortos

Algoritmo de Dijkstra

El algoritmo de Dijkstra es el algoritmo para calcular caminos mas cortos en grafos más usado. Fue concebido por el científico de la computación Edsger Dijkstra en 1956. El algoritmo es iterativo, en cada iteración se actualizan las distancias entre un nodo inicial (definido por el usuario) y cada uno de los nodos restantes del grafo. En principio a cada nodo se le asigna un valor que representa la distancia al **nodo inicial** y el algoritmo busca disminuir este valor. Los pasos del algoritmo son:

- 1 Marcar a todos los nodos como **no visitados**, crear un conjunto de nodos no visitados llamado **el conjunto no visitado**
- 2 Crear una tabla de nodos predecesores llamado el **tabla de predecesores** que contenga a los predecesores de cada nodo. El predecesor de un nodo es el nodo previo en un camino dado.
- 3 Asignarle al nodo inicial una distancia de 0 (la distancia entre un nodo y si mismo es cero) y al resto de los nodos una distancia de infinito.
- 4 Para el nodo actual (en la primera iteración es el **nodo inicial**) calcular las distancias tentativas a cada uno de sus vecinos (nodos directamente conectados) que estén marcados como **no visitados**. Las distancias tentativas se calculan sumando la distancia del nodo actual (en el caso del nodo inicial seria una distancia de cero) y la distancia asociada a la arista que lo une con el nodo vecino.
- 5 Para cada vecino no visitado comparar su distancia tentativa con su distancia asignada, si la distancia tentativa es menor, reemplazar la distancia asignada por la distancia tentativa y marcar en la **tabla de predecesores** al nodo actual como el nodo predecesor de este vecino no visitado.
- 6 Sacar al nodo actual del **conjunto no visitado** (nunca se lo volverá a visitar).

7 Elegir de entre los nodos **no visitados** al que tenga la menor distancia asignada como el nuevo nodo actual.

8 Repetir los pasos **4-7** hasta que el **conjunto no visitado** quede vacío.

Al finalizar el algoritmo, cada nodo tiene asignada la distancia hacia el **nodo inicial** y se puede reconstruir el camino mas corto con la **tabla de predecesores**. En el caso de que la distancia asignada a un nodo sea infinito, quiere decir que ese nodo no puede ser accedido desde el **nodo inicial**, es decir, no existe ningún camino que los una. En la figura 2.2 se muestra una representación grafica del algoritmo en un grafo pesado no dirigido.

Algoritmo de Yen

El algoritmo de Yen es un algoritmo utilizado para obtener los K caminos sin bucles mas cortos que existen entre un **nodo fuente** y un **nodo objetivo**. Fue publicado por Jin Y. Yen en 1971. Utiliza algún algoritmo para obtener el camino mas corto a partir del cual busca las $K - 1$ desviaciones mas cortas. Para explicar como funciona el algoritmo, primero es necesario establecer algunas definiciones:

N: Número de nodos que contiene el grafo.

(i): Nodo i -ésimo del grafo. El nodo fuente es **(1)** y el nodo objetivo es **(N)**.

d_{ij} : La distancia asociada a la arista a_{ij} , es decir $d_{ij} = f(w_{ij})$.

A^k : $A^k = (1) - (2^k) - (3^k) - \dots - (Q_k^k) - (N)$ es el k -ésimo camino mas corto entre **(1)** y **(N)**, con k entre 1 y K . **(2^k)** es el segundo nodo de A^k , **(3^k)** el tercero y así con el resto de los nodos, siendo **(Q_k^k)** el nodo anterior a **(N)**.

A_i^k : Desviación del camino A^{k-1} en el nodo **i^k**, con i entre 1 y Q_k . No puede haber una desviación en el nodo **(N)**. Los caminos A_i^k y A^{k-1} son iguales hasta el nodo i -ésimo.

R_i^k : Camino raíz de A_i^k que es igual a A^{k-1} hasta el nodo i -ésimo.

S_i^k : Camino espurio de A_i^k que empieza en el nodo i -ésimo y termina en el nodo objetivo **(N)**

Los pasos del algoritmo son:

1 Crear el contenedor **A** que guardará a los k caminos mas cortos y el contenedor **B** que contendrá a los potenciales k -caminos mas cortos.

2 Determinar el camino mas corto **A^1** entre el **nodo fuente** y el **nodo objetivo** usando algún algoritmo de camino mas cortos, por ejemplo, el algoritmo de Dijkstra.

3 Elegir el nodo espurio **(i^{k-1})** y el camino raíz **R_i^k** . Si **R_i^k** es igual a **R_i^j** , eliminar la arista **(i^{j-1}) - ((i+1)^{j-1})** con $j \in [2, k]$. Es decir, si el camino raíz **R_i^k** esta incluido en alguno de los caminos mas cortos ya calculados, se borran las aristas que continuaban en esos caminos y se asegura que el camino espurio **S_i^k** no finalice igual que estos. Así, **A_i^k** no estará ya entre los caminos mas cortos anteriormente obtenidos.

4 Eliminar todos los nodos presentes en **R_i^k** para evitar atravesarlos nuevamente.

5 Calcular **S_i^k** como el camino mas corto entre **(i^{k-1})** y **(N)**. Luego obtener **A_i^k** uniendo **R_i^k** y **S_i^k** y agregarlo al contenedor **B**.

6 Restaurar aristas y nodos borrados en **3** y **4**.

- 7 Repetir [3-6](#) para $i \in [1, Q_k]$.
- 8 Encontrar el camino mas corto en \mathbf{B} , agregarlo a \mathbf{A} y eliminarlo de \mathbf{B} . Si \mathbf{B} esta vacío, terminar el algoritmo.
- 9 Repetir pasos [3-8](#) para $k \in [2, K]$.

Finalmente, en \mathbf{A} estarán los K caminos mas cortos entre el **nodo fuente** y el **nodo objetivo**. Si la cantidad de caminos es menor que K , el algoritmo devolverá todos los caminos que haya disponibles.

2.3. Medidas de comunicación en redes

Ya introducidas las redes, su representación y los diferentes enfoques que existen sobre comunicación, surge la pregunta, ¿de que forma podemos medir la comunicación en redes? Esta es una respuesta abierta y posiblemente las respuestas sean incontables, sin embargo, en esta sección, se introducirán algunas (contables) medidas que permiten capturar en cierta forma la comunicación de una red.

Notación

N : Conjunto de nodos del grafo, $|N|$ es el número de nodos, también conocido como tamaño del grafo.

Probabilidad de un camino

La probabilidad asociada a un camino ρ es una medida intrínsecamente asociada a un modelo de comunicación por difusión ya que refleja la probabilidad de que un *random walker* tome el camino en cuestión. Viene dada por

$$P(\pi) = \prod_{i=0}^{|\rho|} \frac{\rho^i}{s_{\Omega^i}}, \quad (2.4)$$

donde ρ^i representa la i -ésima arista y s_{Ω^i} , definida en [2.2](#), la fuerza del i -ésimo nodo del camino π . Cabe destacar que el numero de aristas del camino $|\rho|$ es menor al número de nodos, por lo que [2.4](#) excluye al último nodo del camino, el nodo donde termina y es por esto que la probabilidad de un camino entonces depende del orden en el que se recorre el mismo.

2.3.1. Eficiencia de Ruteo

La eficiencia de ruteo es una medida que sirve para capturar que tan integrada esta una red a través de los caminos mas cortos. Se calcula de la forma:

$$E_R = \frac{1}{|N|(|N|-1)} \sum_{i \neq j \in N} \frac{1}{l_{ij}}, \quad (2.5)$$

donde l_{ij} es la longitud del camino mas corto entre el nodo i y el nodo j , definida en la ecuación [2.3](#). Es decir que E_R es el valor medio de la inversa de la distancia entre todos los pares de nodos. Cuanto mayor sea E_R , mas cerca se encuentran los nodos de la red entre si.

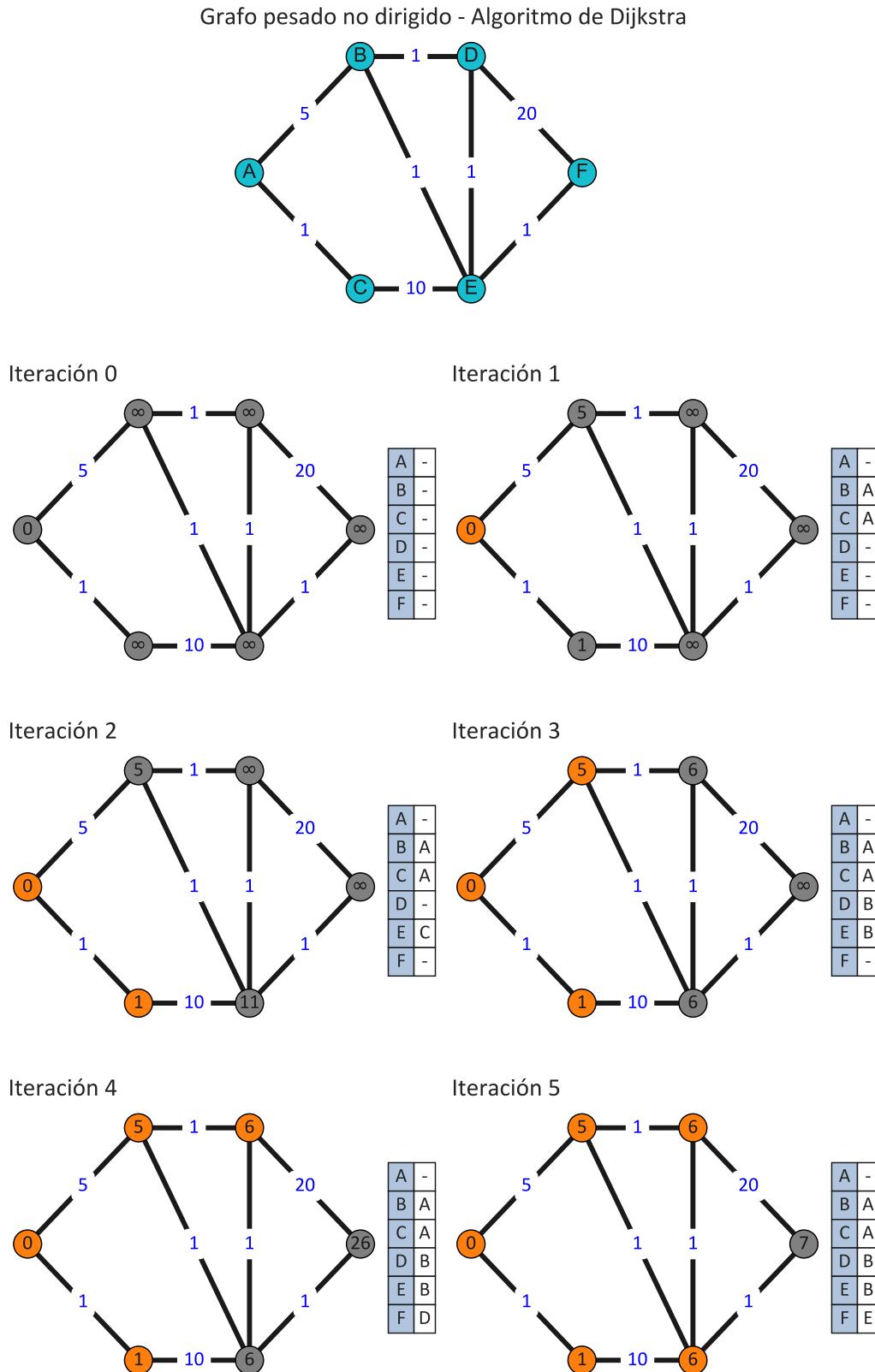


Figura 2.2: Algoritmo de Dijkstra para calcular el camino mas corto entre pares de nodos aplicado en un grafo no dirigido y pesado de 6 nodos. El nodo inicial es el nodo A, cada arista esta marcada con su distancia asociada. Los nodos de color naranja son los nodos visitados y los grises los no visitados, cada nodo esta marcado con su distancia hacia el nodo inicial. En cada iteración se actualiza la distancia asignada de cada vecino del nodo actual y también tabla de predecesores, ubicada a la derecha.

2.3.2. Eficiencia de Difusión

En un contexto de comunicación por difusión, las señales no viajan a través de los caminos mas cortos, si no que siguen una lógica de caminata aleatoria. Entonces, para medir que tan eficientemente conectada esta una red, no sirve una medida basada en caminos mas cortos. Goñi et al (2013) [6] propusieron una medida de la eficiencia de comunicación en un contexto de difusión a la que llamaron eficiencia de difusión y se calcula de la forma:

$$E_D = \frac{1}{|N|(|N|-1)} \sum_{i \neq j \in N} \frac{1}{\langle X_{ij} \rangle}, \quad (2.6)$$

donde $\langle X_{ij} \rangle$ es el numero medio de aristas que un *random walker* tiene que atravesar para llegar desde el nodo i hacia el nodo j . Es decir, cuanto menos aristas se atraviesen en media en una caminata aleatoria entre los nodos de la red, mas eficientemente integrada estará según E_D . Para poder calcular $\langle X_{ij} \rangle$ es necesario hacer uso de la matriz U definida:

$$U = W^T S^{-1} \quad \wedge \quad S = \begin{bmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_{|N|} \end{bmatrix}, \quad (2.7)$$

donde W es la matriz de pesos del grafo y s_i es el grado pesado de un nodo definido en 2.2. El elemento $u_{ji} = w_{ij}/s_j$ se define como la probabilidad de que un caminante aleatorio ubicado en el nodo i se mueva hacia el nodo j . Se puede ver que U^H expresa las probabilidades de que un caminante aleatorio llegue de un nodo hacia otro en H pasos. Es decir, $[U^H]_{ji}$ expresa la probabilidad de que una señal que sigue una caminata aleatoria vaya desde el nodo i hacia el nodo j atravesando exactamente H aristas. Definiendo U_j como la matriz U con todos los elementos de la fila j -ésma puestos a cero, es decir $[U_j]_{ji} = 0 \forall i \in N$, se puede calcular:

$$\langle X_{ij} \rangle = \sum_{t=0}^{\infty} t P(X_{ij} = t) \stackrel{(a)}{=} \sum_{t=0}^{\infty} P(X_{ij} > t) \stackrel{(b)}{=} \sum_{t=0}^{\infty} \sum_{n=1}^{|N|} [U_j^t]_{ni} = \sum_{n=1}^{|N|} [(1 - U_j)^{-1}]_{ni} \quad (2.8)$$

La igualdad (a) viene por la propiedad de que el valor medio de una variable aleatoria discreta positiva es igual a la suma de la función cumulativa complementaria. Es necesario recordar que una señal puede atravesar infinitas aristas en su camino del nodo i al nodo j , por eso la sumatoria va hasta infinito. La igualdad (b) viene por el hecho de que la matriz U_j^t representa las probabilidades de que un mensaje que sigue una caminata aleatoria atraviese exactamente t aristas en su camino entre pares de nodos de la red sin la opción de ir al nodo j . El elemento $[U_j^t]_{ni}$ es la probabilidad de que un caminante aleatorio llegue desde el nodo i hacia el nodo n en exactamente t pasos sin pasar por el nodo j en su camino. Entonces $\sum_{n=1}^{|N|} [U_j^t]_{ni}$ representa la probabilidad de que un mensaje salga del nodo i , no haya pasado por el nodo j y este en cualquier otro nodo diferente a j después de atravesar t aristas en su caminata aleatoria, que es igual a la probabilidad de que no haya llegado al nodo j después de t aristas.

2.3.3. Comunicabilidad

La comunicabilidad es una medida propuesta por Estrada y Hatano (2008) [7] que esta enmarcada en un contexto de flujo de información por difusión. Esta medida tiene en cuenta todos los caminos

posibles entre pares de nodos y se define

$$C_{ij} = \sum_{n=0}^{\infty} \frac{[A^n]_{ij}}{n!} = [e^A]_{ij}, \quad (2.9)$$

donde A es la matriz de adyacencia (binaria) del grafo. Esta medida se basa en que $[A^n]_{ij}$ es el número total de caminos entre el nodo i y el nodo j compuestos por n aristas. La comunicabilidad entonces, es la suma pesada del numero total de caminos entre pares de nodos, con un peso inversamente proporcional al factorial del numero de aristas que componen los componen. Crofts y Higham (2009) [8] generalizaron la medida para ser utilizada en la matriz de pesos W de la forma

$$C_{ij} = \left[\exp \left(S^{-1/2} W S^{-1/2} \right) \right]_{ij} \quad (2.10)$$

donde se usa la matriz S , definida en 2.7, como normalización para regular la influencia de nodos con alta fuerza. En este caso la comunicabilidad ya no es la suma pesada de la cantidad de caminos entre dos nodos, si no que esta relacionada a las probabilidades de los caminos existentes. La comunicabilidad pesada fue usada por Crofts et al [4] para encontrar cambios en la conectividad de conectomas estructurales de pacientes que sufrieron un accidente cerebrovascular respecto a pacientes sanos.

2.3.4. Distancia de K-caminos

En la sección 2.2.1 se introdujo la longitud del camino mas corto como la distancia topológica entre dos nodos de una red. Sin embargo, esta medida no tiene en cuenta todas las posibilidades de flujo de información paralela que presenta una red y que son explotadas en un modelo de difusión. Es por esto que Avena-Koenisberger et al [9] propusieron una extensión de esta medida para agregar la contribución de los K caminos mas cortos entre un par de nodos. Se define \hat{D}_{Kji} de la forma

$$\hat{D}_{Kij} = \sum_{k=1}^K \hat{P}_K(\pi_{ij}^k) l_{ij}^k, \quad (2.11)$$

donde π_{ij}^k es el k -ésimo camino mas corto entre los nodos i y j y l_{ij}^k es la longitud de ese camino. $\hat{P}_K(\rho_{ij}^k)$ es la probabilidad normalizada del k -ésimo camino, definida como

$$\hat{P}_K(\pi_{ij}^k) = \frac{P(\pi_{ij}^k)}{\sum_{k'=1}^K P(\pi_{ij}^{k'})}, \quad (2.12)$$

de forma que $\sum_{k=1}^K \hat{P}_K(\pi_{ij}^k) = 1$ y la distancia \hat{D}_K sea la media pesada de las distancias de todos los K caminos mas cortos, donde los pesos representan la probabilidad con la que se recorre cada uno. Para redes no dirigidas π_{ij}^k esta compuesto por los mismos nodos y aristas que π_{ji}^k , sin embargo, como se menciono al introducir la probabilidad de un camino, $P(\pi_{ij}^k)$ no es necesariamente igual a $P(\pi_{ji}^k)$ ya que la secuencia en que se recorren estos nodos y aristas es diferente. De esta forma \hat{D}_{Kij} puede ser diferente a \hat{D}_{Kji} , por eso se simetriza y se define la distancia de K-caminos entre un par de nodos como

$$D_{Kij} = \frac{\hat{D}_{Kij} + \hat{D}_{Kji}}{2} = D_{Kji}. \quad (2.13)$$

2.3.5. Centralidad de nodos y aristas

La centralidad es una medida que captura la relevancia de cada nodo y arista en los caminos mas cortos de una red. La centralidad de nodos NC y la centralidad de aristas EC están definidas originalmente sobre caminos mas cortos, sin embargo, Avena-Koenisberger et al. [9] la extendieron para incluir los K caminos mas cortos entre pares de nodos de la forma

$$NC_i = \frac{1}{K(|N| - 1)(|N| - 2)} \sum_{\substack{g \neq i \\ h \neq i \\ k \leq K}} \sigma(\Omega_{gh}^k, i), \quad (2.14)$$

$$EC_{ij} = \frac{1}{K|N|(|N| - 2)} \sum_{\substack{g \neq h \\ k \leq K}} \sigma(\rho_{gh}^k, w_{ij}), \quad (2.15)$$

$$\sigma(x, y) = \begin{cases} 1 & \text{si } y \in x \\ 0 & \text{en otro caso.} \end{cases} \quad (2.16)$$

Así, la centralidad mide la cantidad de veces que un nodo o una arista dada es atravesada por uno de los K caminos mas cortos de forma normalizada, donde una centralidad de 1 significaría que todos los caminos pasan por ese nodo/arista. Conocer EC y NC da información sobre la resiliencia de una red. Si hay pocos nodos o aristas muy centrales, quitarlos o afectarlos sería altamente perjudicial para la comunicación de la red, ya que se verían afectados muchas de las vías de información (caminos). En cambio, si hay valores de EC y NC distribuidos uniformemente, retirar algunos de los elementos de la red no afectaría mucho a la comunicación de la misma.

2.3.6. *Search information*

2.3.7. Matriz de acoplamiento

La matriz de acoplamiento G captura

Apéndice A

Apéndice I

Bibliografía

- [1] Mori, S. Three-dimensional tract reconstruction. En: Introduction to Diffusion Tensor Imaging, cap. 9. ELSEVIER, 2007. 5
- [2] Smith, R. E., Tournier, J.-D., Calamante, F., Connelly, A. Sift2: Enabling dense quantitative assessment of brain white matter 2 connectivity using streamlines tractography. *NeuroImage*, **119**, 338–351, 2015. 4
- [3] Goñi, J., et al. Resting-brain functional connectivity predicted by analytic measures of network communication. *PLOS ONE*, **8** (3), 1–10, 2014. 4
- [4] Crofts, J., et al. Network analysis detects changes in the contralesional hemisphere following stroke. *NeuroImage*, **54**, 161–169, 2011. 4, 15
- [5] Planchuelo-Gómez, Á., García-Azorín, D., L Guerrero, Á., Aja-Fernández, S., Rodríguez, M., de Luis-García, R. Structural connectivity alterations in chronic and episodic migraine: A diffusion magnetic resonance imaging connectomics study. *Cephalgia*, **40** (4), 367–383, 2019. 6
- [6] Goñi, J., Avena-Koenigsberger, A., Velez de Mendizabal, N., van den Heuvel, M. P., Betzel, R. F., Sporns, O. Exploring the morphospace of communication efficiency in complex networks. *PNAS*, **111** (2), 833–838, 2013. 14
- [7] Estrada, E., Hatano, N. Communicability in complex networks. *Physical review. E, Statistical, nonlinear, and soft matter physics*, **77**, 036111, 04 2008. 14
- [8] Crofts, J., Higham, D. A weighted communicability measure applied to complex brain networks. *Journal of the Royal Society, Interface / the Royal Society*, **6**, 411–4, 02 2009. 15
- [9] Avena-Koenigsberger, A., Misic, B., Hawkins, R., Griffa, A., Hagmann, P., Goñi, J., et al. Path ensembles and a tradeoff between communication efficiency and resilience in the human connectome. *Brain structure & function*, **222**, 01 2017. 15, 16

Publicaciones asociadas

1. Mi primer aviso en la revista **ABC**, 1996
2. Mi segunda publicación en la revista **ABC**, 1997

Agradecimientos

A todos los que se lo merecen, por merecerlo

