Grafos (1/3): Tipos de grafos / redes y algunas propiedades locales.

Bibliografía

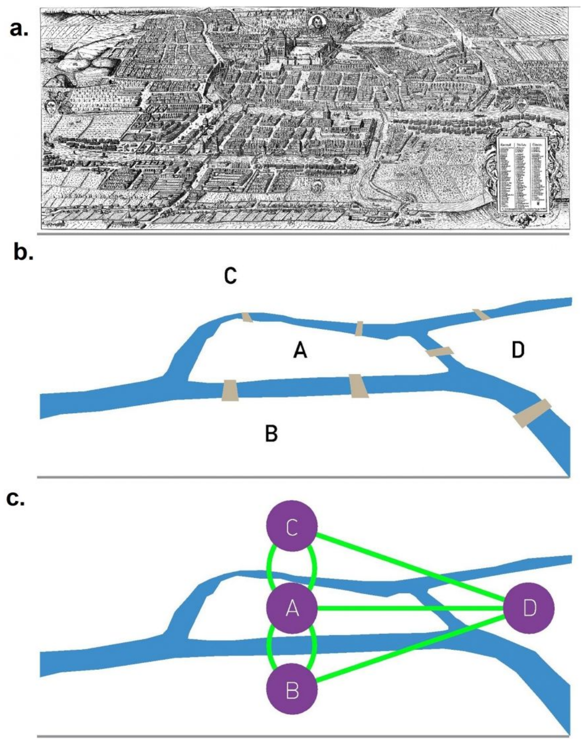
Barabási, A. L. (**2016**). *Network science.* Cambridge university press. http://networksciencebook.com/

Kolaczyk, E. D., & Csárdi, G. (**2014**). *Statistical analysis of network data with R* (Vol. 65). New York: Springer.

Sporns, O. (**2010**). *Networks of the Brain*. MIT press.

Kolaczyk, E. D. (**2009**). *Statistical analysis of network data* (Vol. 65). New York: Springer.

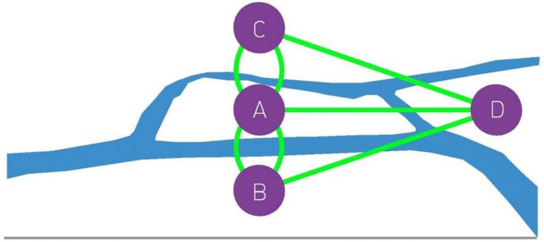
Grafos

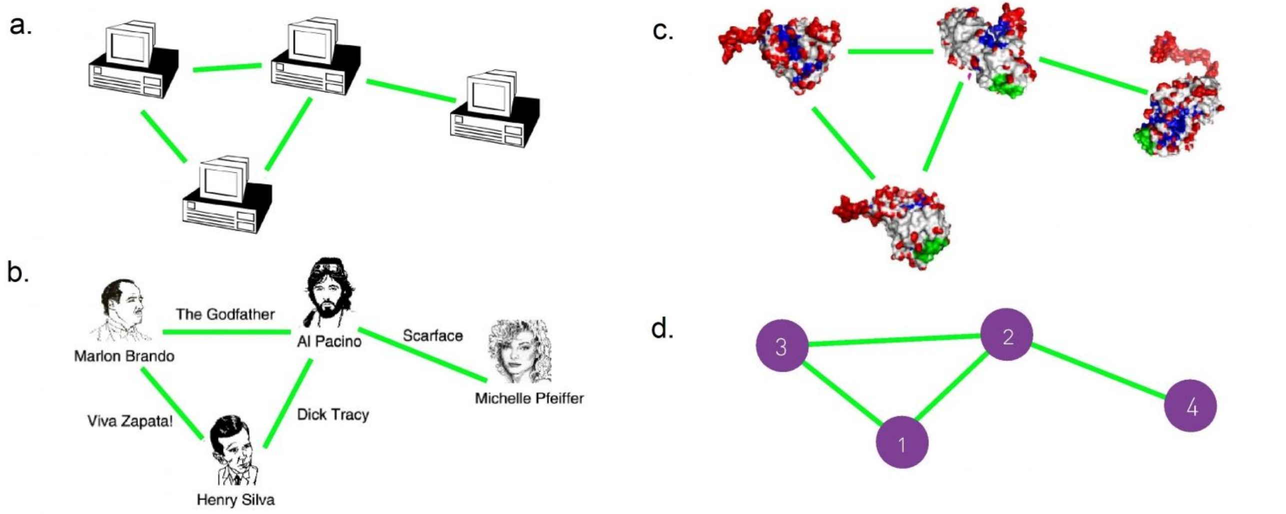


Barabási, A. L. (**2016**). *Network science.* Cambridge university press.

Grafos

c.





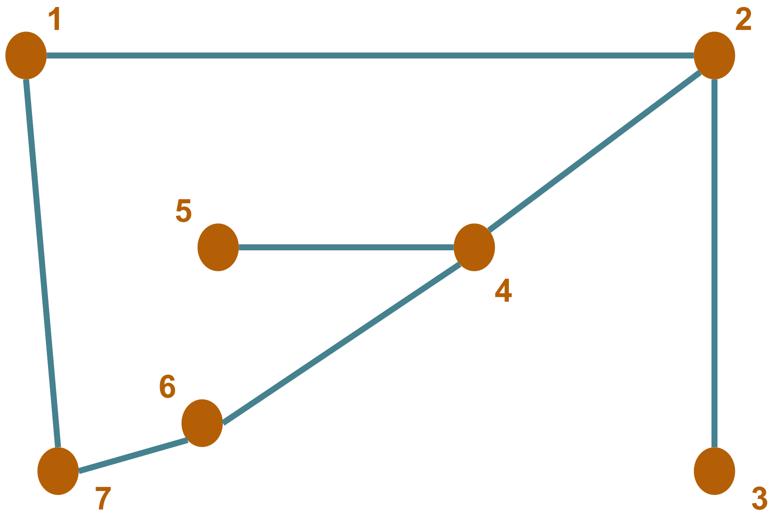
Barabási, A. L. (**2016**). *Network science.* Cambridge university press.

Definiciones y nomenclatura

Grafo : **G = (V,E)**

V: Conjunto de nodos o vértices del grafo

E: Conjunto de aristas



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

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

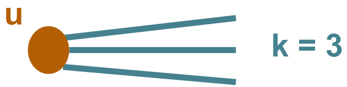
Definiciones y nomenclatura

{u, v} es una arista **e** que conecta los nodos **u** y **v**

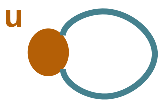
u y **v** son <u>incidentes</u> con **e**

u y **v** son <u>adyacentes</u> o <u>vecinos</u> entre sí

Los vértices o nodos incidentes sobre una arista son los <u>nodos terminales</u> de ella



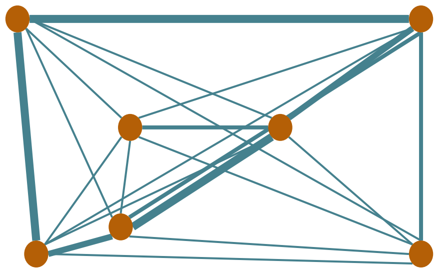
El <u>grado</u> (**k**) de un nodo **u** es el número de aristas que tienen a **u** como nodo terminal.



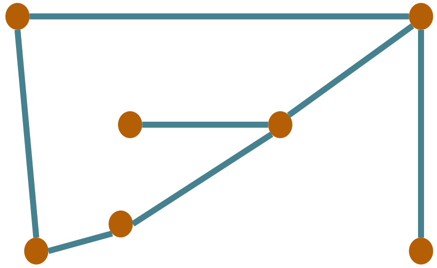
Una arista en la que los dos nodos terminales corresponden al mismo nodo es un <u>loop</u> o <u>lazo</u>.

Tipos de grafos básicos

Grafos ponderados: grafos donde las conexiones tienen un peso.

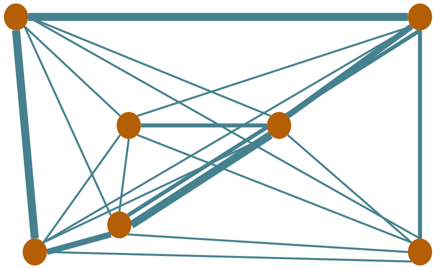


Grafos binarios: grafos donde las conexiones son binarias, es decir que dos vértices, están o no están conectados.

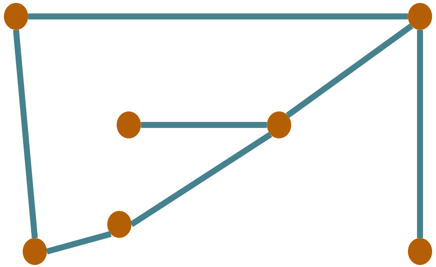


Tipos de grafos básicos

Grafos ponderados: grafos donde las conexiones tienen un peso.



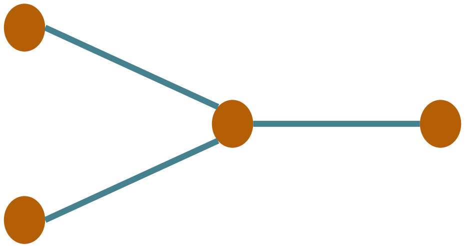
Grafos binarios: grafos donde las conexiones son binarias, es decir que dos vértices, están o no están conectados.



Muchas veces se aplica un criterio para quedarse con el grafo binario, es decir un *umbral fijo* o elegir las *N aristas de mayor peso*.

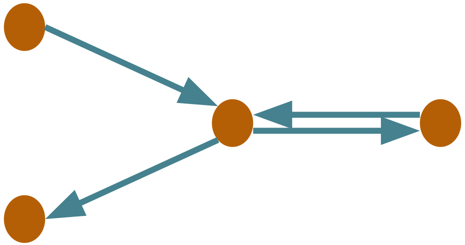
Tipos de grafos básicos

Grafos no dirigidos: grafos donde las conexiones NO tienen un sentido.



Una arista no dirigida tiene nodos terminales.

Grafos dirigidos: grafos donde las conexiones tienen un sentido.

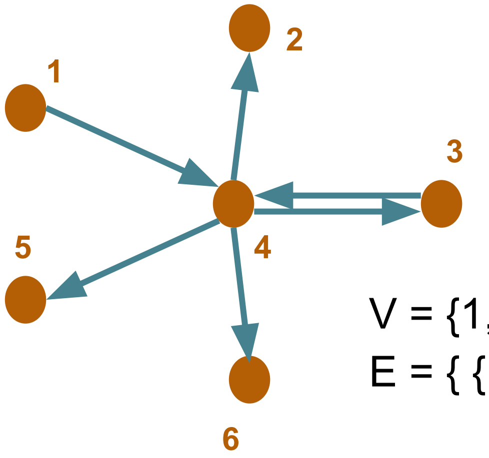


Una arista dirigida tiene un nodo fuente y otro nodo destino.
Un nodo de un grafo dirigido tiene un grado de entrada y otro grado de salida.

Grafos mixtos: grafos donde algunas conexiones tiene sentido y otras no.

Más definiciones y nomenclatura

Grafos dirigidos: grafos donde las conexiones tienen un sentido.

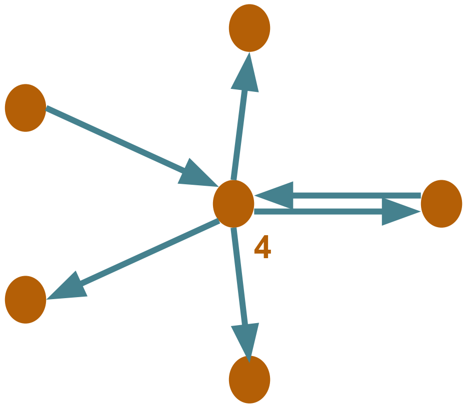


V = {1,2,3,4,5,6}

E = { {1,4}, {4,2}, {3,4}, {4,3}, {4,5}, {4,6}}

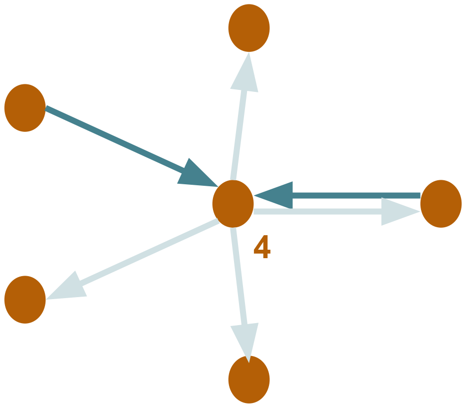
Más definiciones y nomenclatura

Grafos dirigidos: grafos donde las conexiones tienen un sentido.



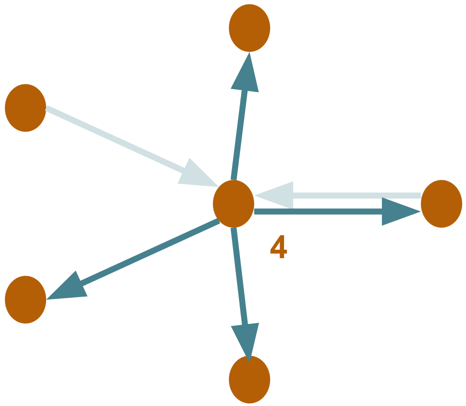
Grado (k_4) = 6

{ {1,4}, {4,2}, {3,4}, {4,3}, {4,5}, {4,6}}



Grado de entrada ($k_{in,4}$) = 2

{ {1,4}, {3,4}}

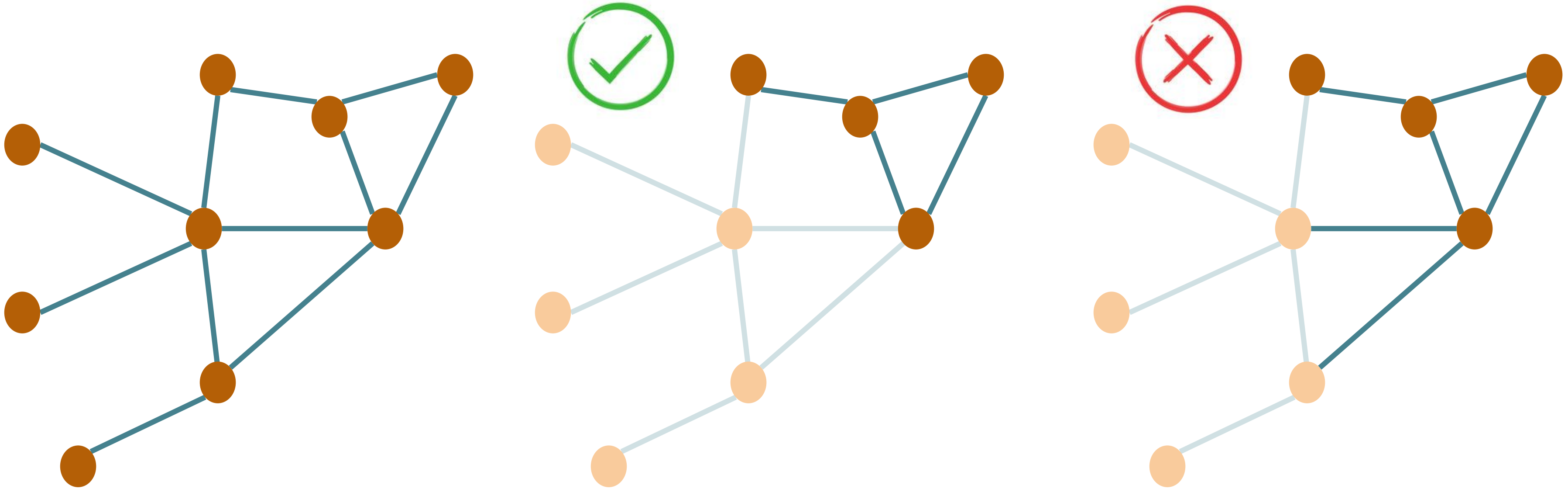


Grado de salida ($k_{out,4}$) = 4

{ {4,2}, {4,3}, {4,5}, {4,6}}

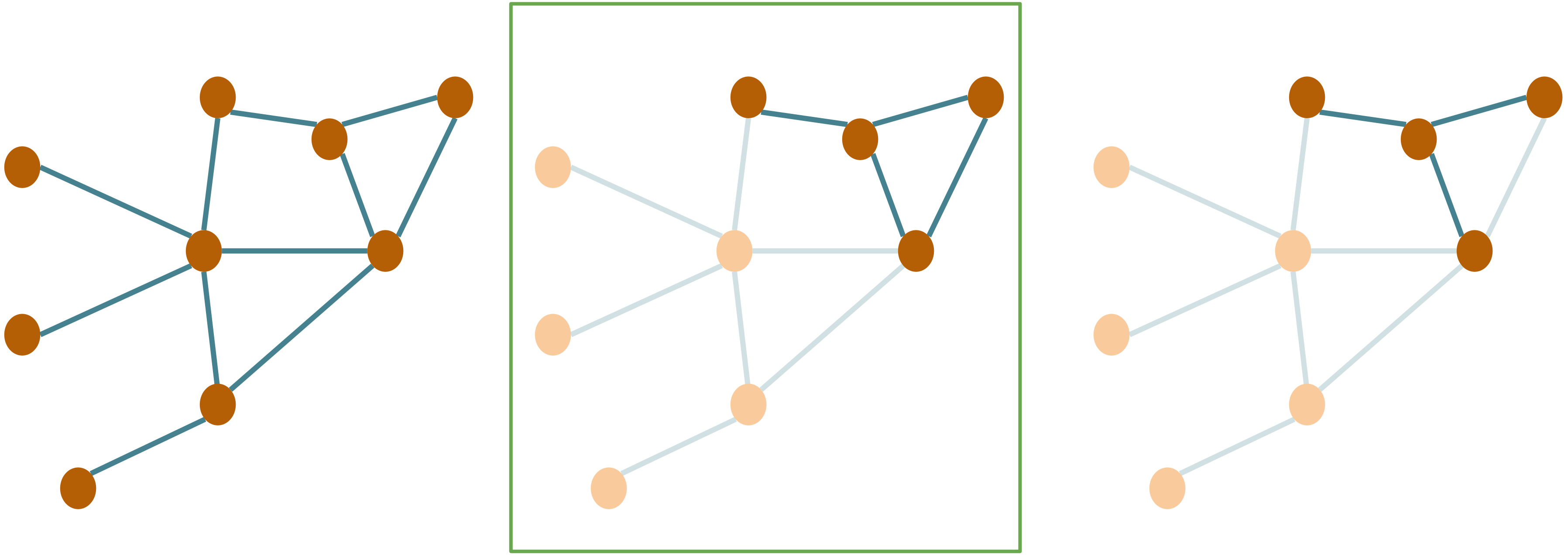
Más definiciones y nomenclatura

Un **subgrafo** *G'=(V',E')* del *G=(V,E)* es un grafo donde *V'* es un subconjunto de *V* y *E'* es un subconjunto de *E*, donde *E'* sólo contiene aristas cuyos nodos terminales se encuentran en *V'*.

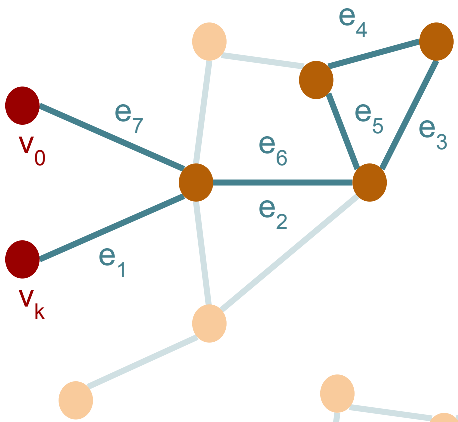


Más definiciones y nomenclatura

Si E' tiene todas las aristas de E que conectan los nodos de V', entonces G' es un **subgrafo inducido** de G=(V,E).

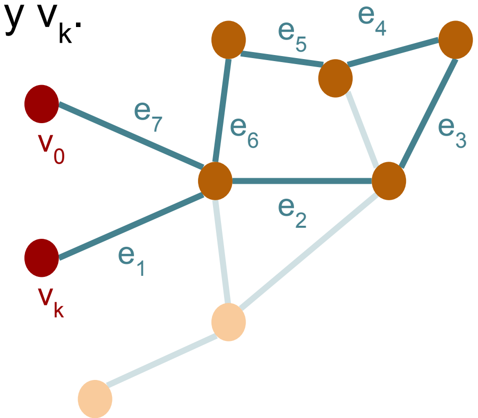


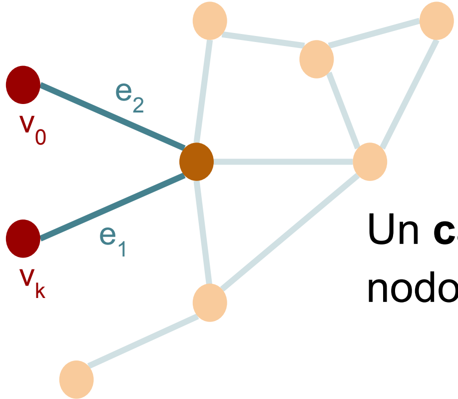
Recorridos y caminos



Un **recorrido (walk)** es una secuencia (e_1, e_2, e_3, … , e_{k-1} , e_k) donde los nodos terminales de e_2 son a su vez terminales de e_1 y e_3. Dicho recorrido tiene siempre dos nodos terminales v_0 y v_k.

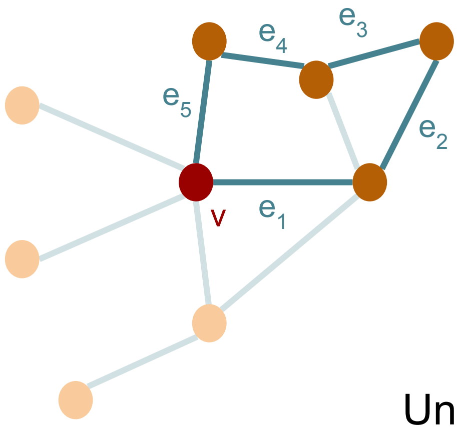
Un **camino (path)** es un recorrido en el que las aristas son todas diferentes.





Un **camino simple (simple path)** es un camino que tiene todos los nodos distintos.

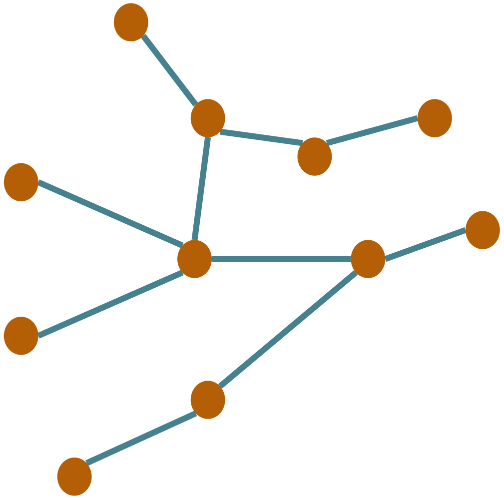
Recorridos y caminos



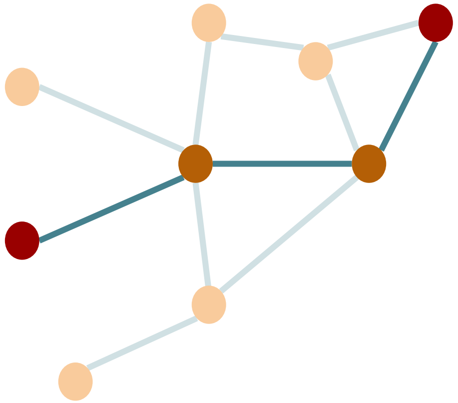
Un **ciclo** es un camino que regresa a su punto de partida.

Un **grafo acíclico** es un grafo que no tiene ciclos.

((ÁRBOLES))



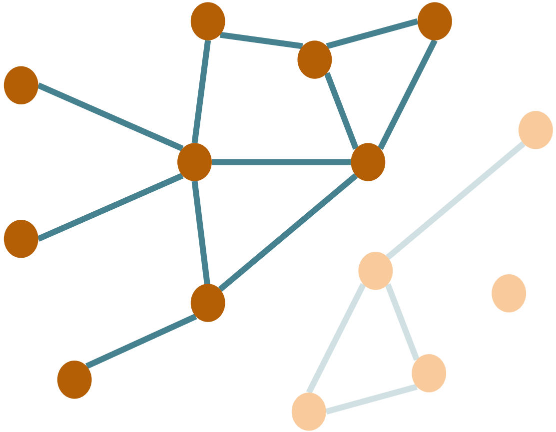
Recorridos y caminos



Dos nodos están **conectados** si existe un camino entre ellos.

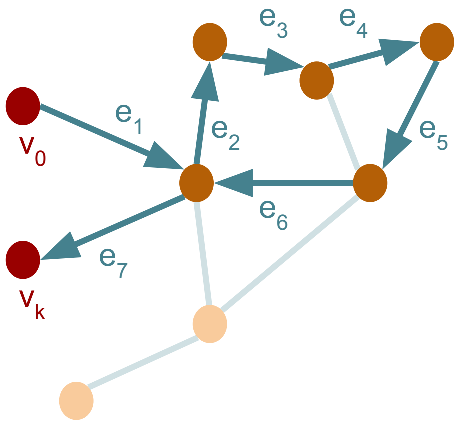
Si cualquier par de vértices del grafo está conectado, se dice que el grafo es **conexo**.

Un **componente conexo** es el máximo subgrafo conexo de de G.



Recorridos y caminos - Grafos dirigidos

En un **grafo dirigido** las definiciones de recorrido y camino son similares pero hay que tener en cuenta el sentido:



Un **recorrido** es una secuencia donde los nodos la arista e_i tiene nodo fuente v_{i-1} y nodo destino v_i. Dicho recorrido va a su vez de un nodo inicial v_0 a un nodo terminal v_k.

Recorridos y caminos - Grafos dirigidos

Si la conexión entre v_0 y v_k se concreta al considerar el sentido, se dice que los dos nodos están **fuertemente conectados**. Si no, simplemente están **conectados**.

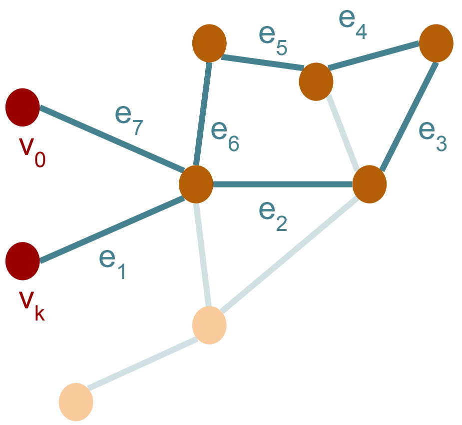
Un grafo está **fuertemente conectado** si cualquier par de nodos está fuertemente conectado.

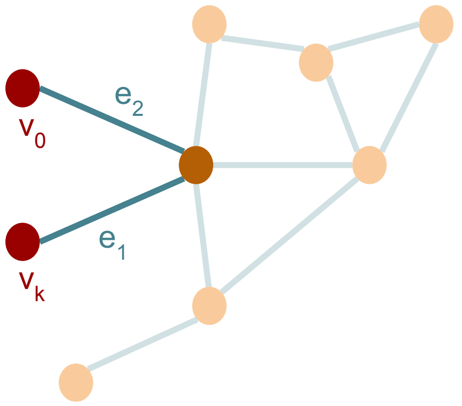
El **componente fuertemente conectado** de G es el máximo subgrafo de G fuertemente conectado.

Longitudes y distancias

La **longitud** de un recorrido o camino está dada por el número de aristas que contienen.

El **camino más corto (shortest path)** entre dos nodos es un camino con longitud mínima. Puede haber más de uno.





Longitudes y distancias

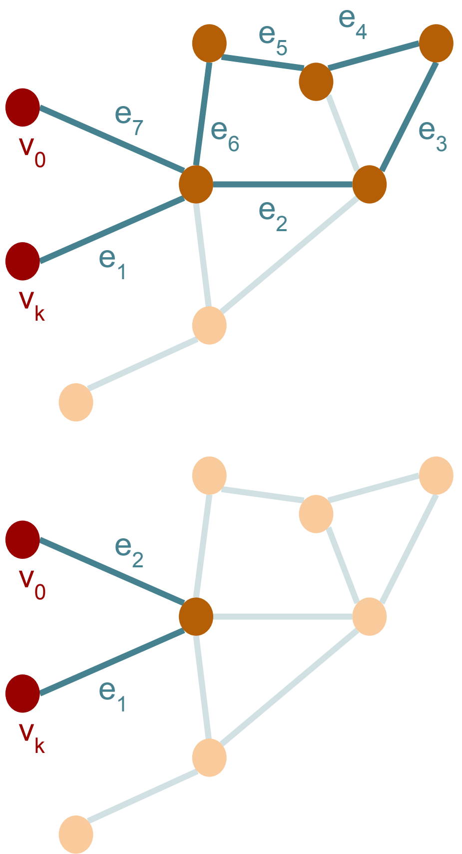
La **distancia** entre dos nodos es la longitud del camino más corto entre ellos, en caso de grafos ponderados es la suma de los pesos.

En grafos dirigidos es común que:

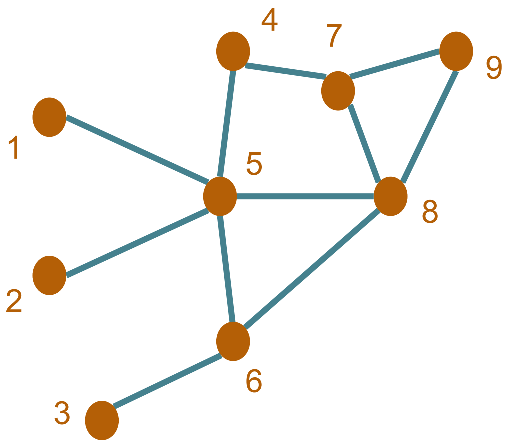
$$\boxed{d_{uv} \neq d_{vu}}$$

En grafos disconexos van a existir uno o más nodos con:

$$\boxed{d_{uv} = \infty}$$

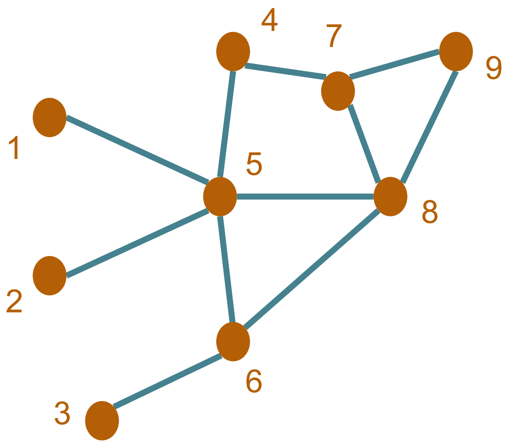


Representación de grafos



Representación de grafos

Matriz de adyacencia ($A_{ij} = \{0,1\}$, $A_{ij} = A_{ji}$)

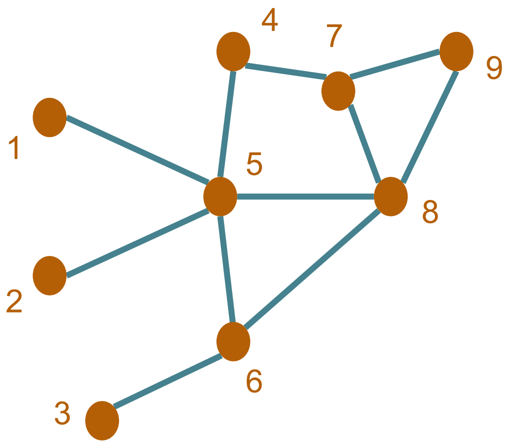


	1	2	3	4	5	6	7	8	9
1	0	0	0	0	1	0	0	0	0
2	0	0	0	0	1	0	0	0	0
3	0	0	0	0	0	1	0	0	0
4	0	0	0	0	1	0	1	0	0
5	1	1	0	1	0	1	0	1	0
6	0	0	1	0	1	0	0	1	0
7	0	0	0	1	0	0	0	1	1
8	0	0	0	0	1	1	1	0	1
9	0	0	0	0	0	0	1	1	0

En un **grafo binario** y **no dirigido**.

Representación de grafos

Matriz de adyacencia (A_{ij})



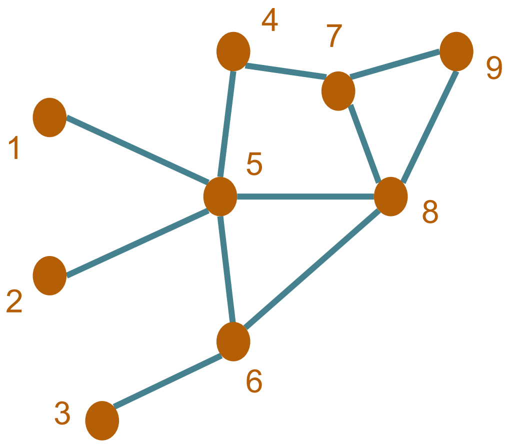
	1	2	3	4	5	6	7	8	9
1	0	0	0	0	1	0	0	0	0
2	0	0	0	0	1	0	0	0	0
3	0	0	0	0	0	1	0	0	0
4	0	0	0	0	1	0	1	0	0
5	1	1	0	1	0	1	0	1	0
6	0	0	1	0	1	0	0	1	0
7	0	0	0	1	0	0	0	1	1
8	0	0	0	0	1	1	1	0	1
9	0	0	0	0	0	0	1	1	0

En un **grafo ponderado** los 0 y 1 cambian por números que representan un peso.

En un **grafo dirigido** la matriz no es simétrica.

Representación de grafos

Lista de adyacencia



L 1 : ({1,5})
L 2 : ({2,5})
L 3 : ({3,6})
L 4 : ({4,5}, {4,7})
L 5 : ({5,1}, {5,2}, {5,4}, {5,6}, {5,8})
L 6 : ({6,3}, {6,5}, {6,8})
L 7 : ({7,4}, {7,8}, {7,9})
L 8 : ({8,5}, {8,6}, {8,7}, {8,9})
L 9 : ({9,7}, {9,8})

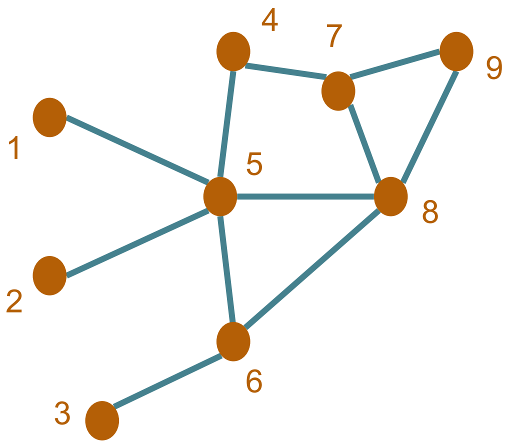
En un grafo **ralo (sparse)**, que se observan frecuentemente:

$$N_e << N_v^2$$

esta representación es más eficiente.

Representación de grafos

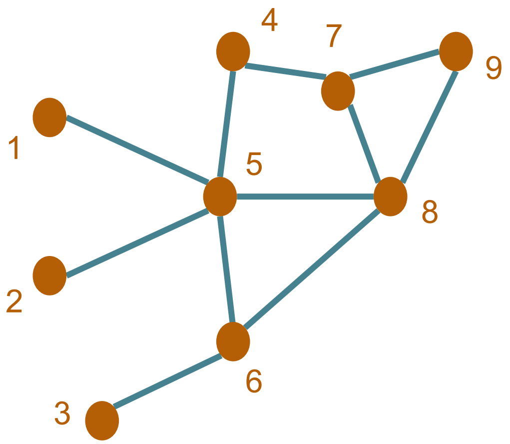
Matriz de distancias



	1	2	3	4	5	6	7	8	9
1	0	2	3	2	1	2	3	2	3
2	2	0	3	2	1	2	3	2	3
3	3	3	0	3	2	1	3	2	3
4	2	2	3	0	1	2	1	2	2
5	1	1	2	1	0	1	2	1	2
6	2	2	1	2	1	0	2	1	2
7	3	3	3	1	2	2	0	1	1
8	2	2	2	2	1	1	1	0	1
9	3	3	3	2	2	2	1	1	0

Propiedades de los vértices. Medidas de centralidad

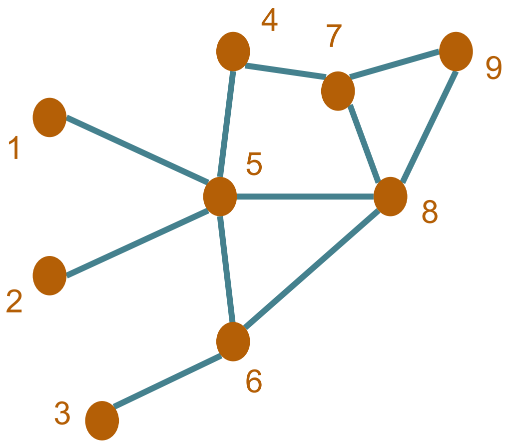
Definiciones



- ➔ De manera abstracta la **centralidad (*C*)** es una función que asigna un valor numérico a cada nodo de una red.
- ➔ El nodo *i* es más central que el nodo *j* si ***C(i) > C(j)***.

Propiedades de los vértices. Medidas de centralidad

Definiciones



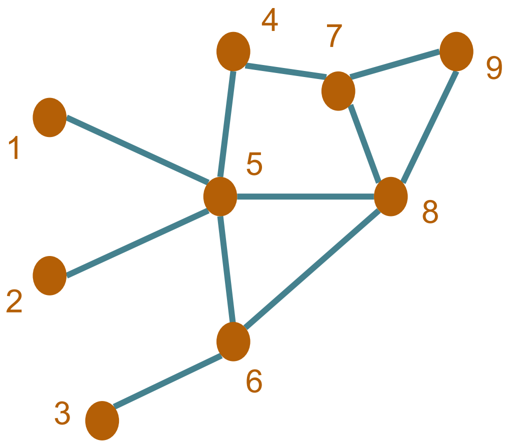
➔ Las medidas de centralidad sirven para caracterizar la conectividad de los nodos de una red.

➔ Los valores de centralidad son comparables sólo dentro de una red.

➔ Algunas medidas de centralidad solo se pueden aplicar a redes conexas.

Propiedades de los vértices. Medidas de centralidad

Ejemplos



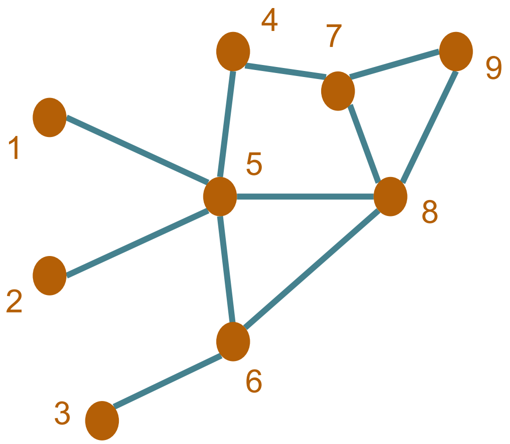
Centralidad de grado: es el **grado** (*k*) del nodo.

También se puede corregir por el número máximo de conexiones *n-1*, para que quede definida en [0 1].

En grafos dirigidos se pueden considerar en forma separada los grados de entrada y de salida, dando una centralidad de entrada y una de salida

Propiedades de los vértices. Medidas de centralidad

Ejemplos



Centralidad de excentricidad (o excentricidad):

$$C_{exc}(i) = \frac{1}{max\{dist(i,j) : j \in V\}}$$

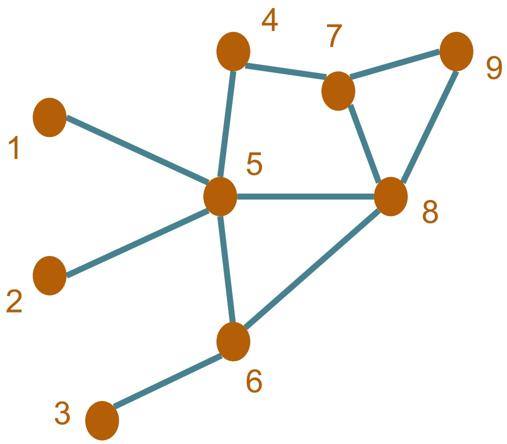
i,*j*: nodos

Se calcula determinando cuál es la máxima distancia entre el nodo *i* y los otros nodos *j*.

Para calcular la excentricidad en redes dirigidas se requiere que éstas estén fuertemente conectadas

Propiedades de los vértices. Medidas de centralidad

Ejemplos



Centralidad de cercanía (closeness):

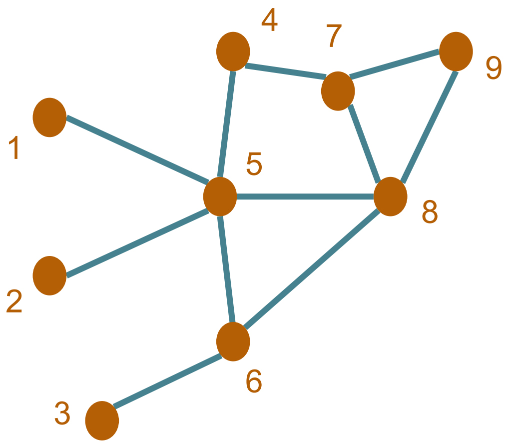
$$C_{cer}(i) = \frac{1}{\sum_{j \in V} dist(i,j)}$$

i,j: nodos

Es una medida basada en la **distancia**.

Propiedades de los vértices. Medidas de centralidad

Ejemplos



Centralidad de intermediación (betweenness):

Siendo *G=(V,E)* un grafo no dirigido con *i,s,t*: nodos

$$C_{bet}(i) = \sum_{s \neq i} \sum_{t \neq i} \delta_{st}(i)$$

donde:

$$\delta_{st}(i) = \frac{\sigma_{st}(i)}{\sigma_{st}}$$

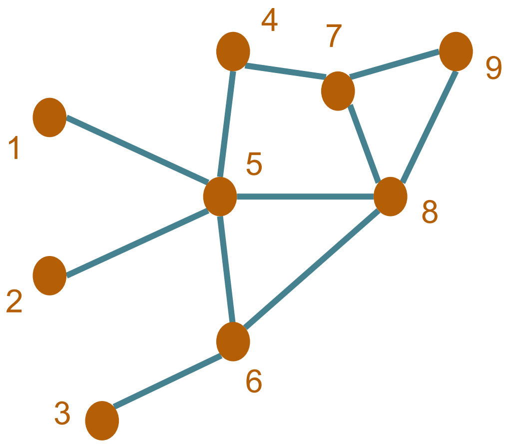
$\sigma_{st}(i)$: **caminos más cortos** entre *s* y *t* con *i* como nodo interno

σ_{st} : **caminos más cortos** entre *s* y *t*.

La medida de intermediación mide la habilidad de un nodo de monitorear las comunicaciones entre otros nodos.

Propiedades de los vértices. Medidas de centralidad

Ejemplos



Centralidad de autovector (eigenvector centrality):

Siendo *G=(V,E)* un grafo no dirigido con *i,s,t*: nodos

$$x_\nu = \frac{1}{\lambda} \sum_{t \in M(\nu)} x_t = \frac{1}{\lambda} \sum_{t \in G} a_{\nu,t} x_t$$

$M(\nu)$: Conjunto de vecinos del nodo ν

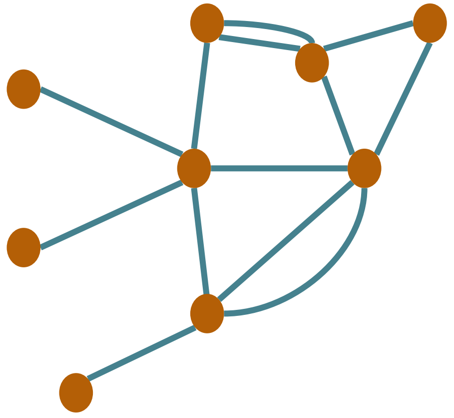
$A = a_{\nu,t}$: Matriz de adyacencia

λ : Máximo autovalor de los autovectores de A ($A\,x = \lambda\,x$)

Esta medida combina la centralidad del vértice ν con las centralidades de sus vecinos directos.

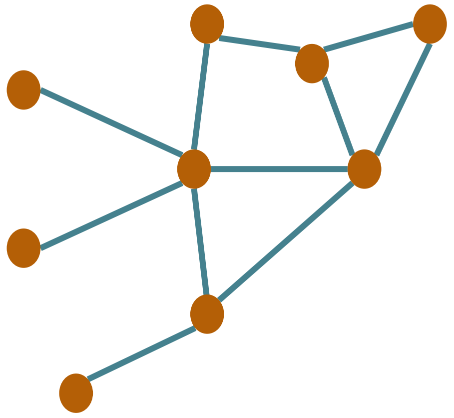
Más tipos de Grafos

Multigrafo



Es un grafo con dos o más aristas múltiples o paralelas que son incidentes sobre el mismo par de vértices. Es un grafo dirigido con las aristas múltiples en la misma dirección.

Grafo simple

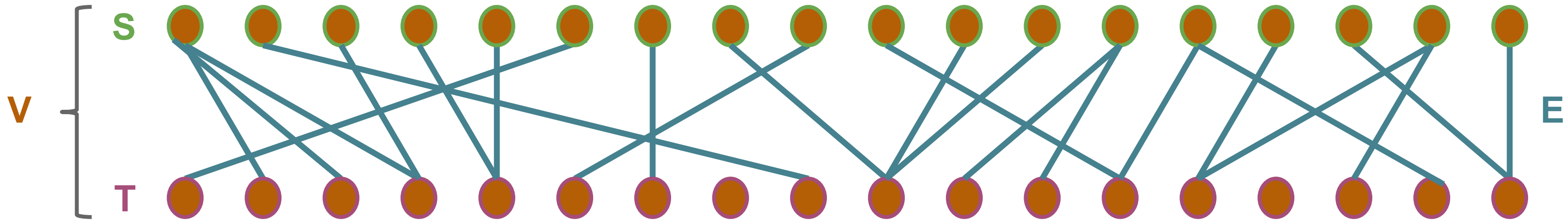


Es un grafo no dirigido, sin loops ni aristas múltiples.

Más tipos de Grafos

Grafo Bipartito

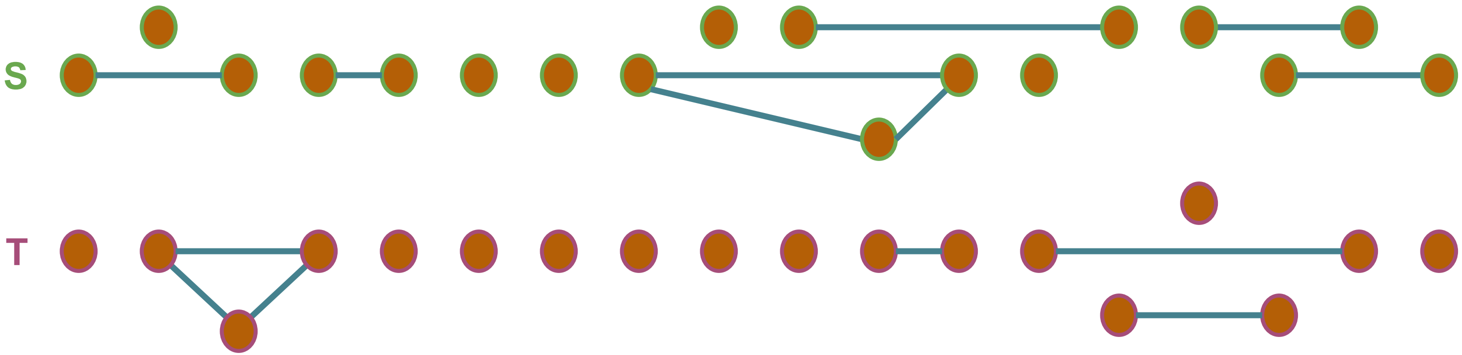
G(V,E) es un **grafo bipartito** si se puede particionar el conjunto de nodos V, en S y T, tal que V = S u T y cada arista en E tiene un nodo en S y otro en T.



Más tipos de Grafos

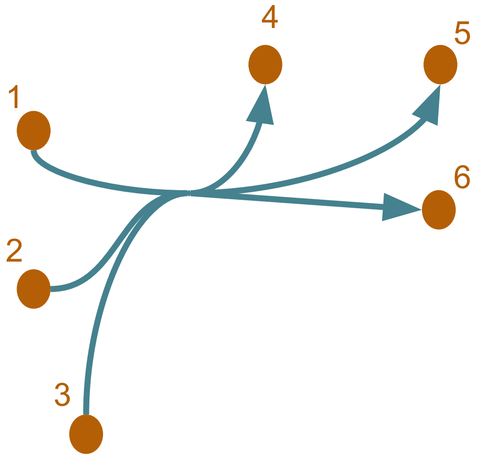
Grafo Bipartito

G(V,E) es un **grafo bipartito** si se puede particionar el conjunto de nodos V, en S y T, tal que V = S u T y cada arista en E tiene un nodo en S y otro en T.



Más tipos de Grafos

Hipergrafo

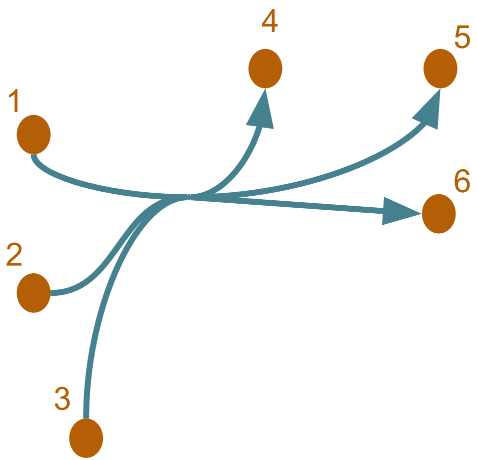


Un hipergrafo G(V,E) se compone de un conjunto de nodos y un conjunto de hiperaristas.

Cada hiperarista es un subconjunto de varios elementos de V. Estos grafos pueden ser dirigidos o no.

Más tipos de Grafos

Hipergrafo



Los hipergrafos son comunes en el mundo real pero no se usan mucho en teoría de grafos. Muchos algoritmos para grafos no pueden procesar hipergrafos.

La solución es modelar un hipergrafo con **grafos bipartitos**.

Más tipos de Grafos

Hipergrafo

¿Cómo modelar un hipergrafo con un grafo bipartito?

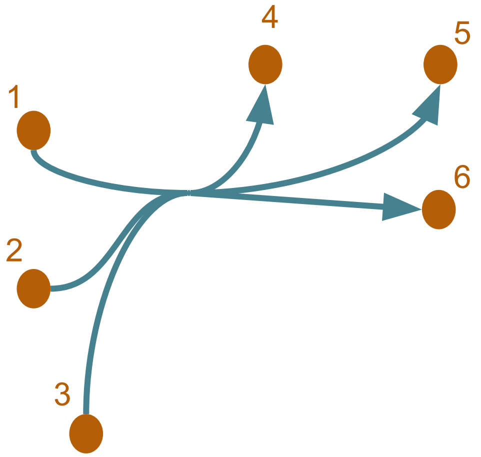
G = (V, E)	hipergrafo
G’= (v’, E’) , V’ = S’ u T’	grafo bipartito

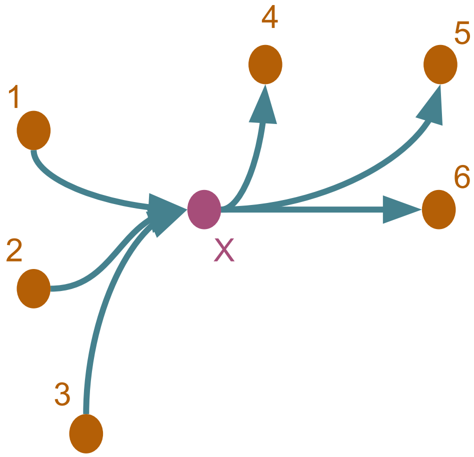
1. Cada nodo en V se representa como un nodo en S’.
2. Cada hiperarista en E se representa con un nodo en T’.
3. Para cada nodo v ∊ V y cada hiperarista e ∊ E incidente con v, se inserta una arista en E’ que conecta un nodo de s ∊ S’ que representa a v y un nodo t ∊ T’, que representa a la hiperarista e.

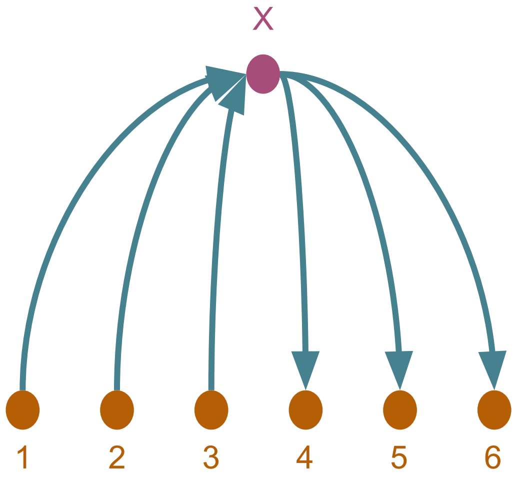
Más tipos de Grafos

Hipergrafo

¿Cómo modelar un hipergrafo con un grafo bipartito?







{ {1,2,3}, {4,5,6}}

{ {1,X}, {2,X}, {3,X}, {X,4}, [X,5}, {X,6}}

Más tipos de Grafos

Árboles

Un árbol es un grafo acíclico, conectado, no dirigido, cuyos nodos de grado 1 se llaman **hojas**, y todos los otros nodos se llaman **nodos internos**.

Un **árbol con raíz** es un árbol $G=(V,E)$ con un nodo distintivo $r \in V$ llamado raíz.

Es común considerar a un árbol con raíz como un grafo dirigido en el que todas las aristas se alejan de la raíz. Se establecen relaciones de padre-hijo entre nodos vecinos.

La **profundidad** de un nodo es la longitud del camino entre la raíz y dicho nodo.

Más tipos de Grafos

Árboles

La **altura** de un árbol es la máxima profundidad que pueda tener un nodo.

Un **árbol binario** es aquel donde cada nodo tiene como máximo grado 3.

Un **árbol abarcativo** T de un grafo no dirigido G= (V,E) es un árbol compuesto por todos los nodos de G y un conjunto mínimo de aristas que conectan a todos los nodos de V.

Grafos (2/3)