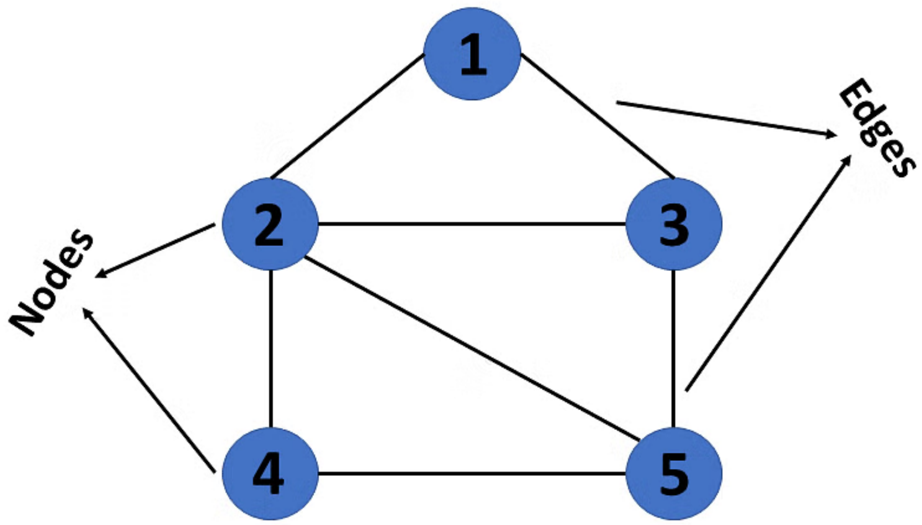
Técnicas y Diseños de Algoritmos

Tomás Agustín Hernández

Grafos

Un grafo es una estructura de datos compuesta por **V un conjunto de vértices (nodos)** y **E un conjunto de aristas que conectan pares de vértices**.
Definimos un grafo como: $G = (V, E)$ donde $E \subseteq VXV$.



Nótese que tiene sentido que el conjunto de aristas E esté formado por un par VXV. De lo contrario no existiría una arista.
Good to Know

- Si dos vértices están conectados por una arista, entonces estos se llaman **adyacentes**.
- La relación entre dos vértices define a $e = (u, v) \in E$. En este caso, decimos que **e** es incidente a **u** y **v**.
- La cantidad de vértices la notamos con la letra **n** y la cantidad de aristas con la letra **m**.

¿Por qué le decimos vértices a los nodos?

Porque los grafos tienen un fundamento mucho más matemático.

Self Loop

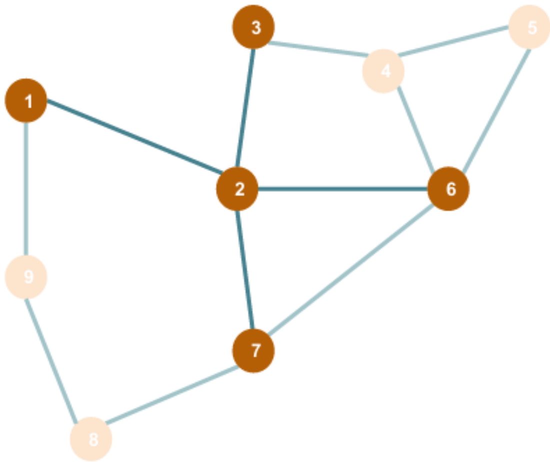
En un grafo, hablamos de self loop cuando un vértice se relaciona consigo mismo.

Cantidad de vertices y cantidad de aristas

Definimos la cantidad de vértices de un grafo como $|V| = n$ y la cantidad de aristas como $|E| = m$.
Good to Know: En un grafo T (árbol) **siempre** sucede que $|V| > |E|$ en una unidad. Lo veremos más adelante, pero está bueno notarlo.

Vecindad de un vértice

Llamamos vecindad de un vértice v a todos los vértices adyacentes de v.
Formalmente, lo definimos como: $N(v) = \{u \in V : (v, u) \in E\}$



Ej.: $N(2) = \{1, 3, 7, 6\}$
Good to Know: Se utiliza la letra N de Neighbor (vecino) en inglés.

Grafo Simple

Un Grafo Simple es un grafo que no tiene más de una arista entre dos mismos nodos. Nótese que aquí no importa la relación entre ellos, sino que, no debería estar repetido el mismo par (sin importar el orden).

- $E = \{(A, B), (A, B)\}$ NO es un Grafo Simple.
- $E = \{(A, B), (B, A)\}$ o $E = \{(B, A), (A, B)\}$ NO es un Grafo Simple.
- $E = \{(A, B)\}$ SÍ es un Grafo Simple.

Grafo no Simple

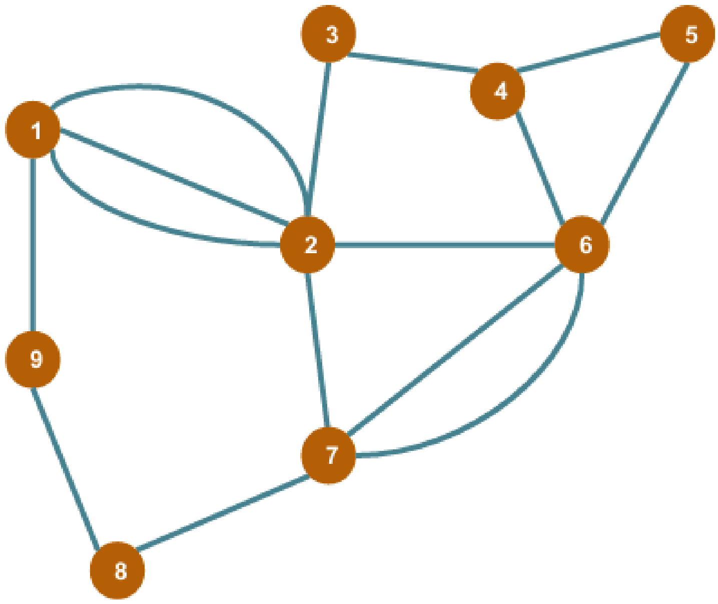
Un Grafo no Simple es un grafo que no tiene ninguna restricción con respecto a la relación entre dos vértices (nodos).

- $E = \{(A, B), (A, B)\}$ es un Grafo no Simple.
- $E = \{(A, B), (B, A)\}$ o $E = \{(B, A), (A, B)\}$ es un Grafo no Simple.
- $E = \{(A, B)\}$ SÍ es un Grafo no Simple.

Good to Know: Los grafos simples, son subconjuntos de grafos no simples. Por ende, si tenemos un grafo no simple que cumple las propiedades de un grafo simple, consideramos que es un **grafo simple** al ser más restrictivo.
En los ejemplos anteriores, entonces, $E = \{(A, B)\}$ sería mejor considerado como Grafo Simple.

Multigrafo

Un Multigrafo es un tipo de **grafo no simple** el cual puede haber más de una arista entre el mismo par de vértices.



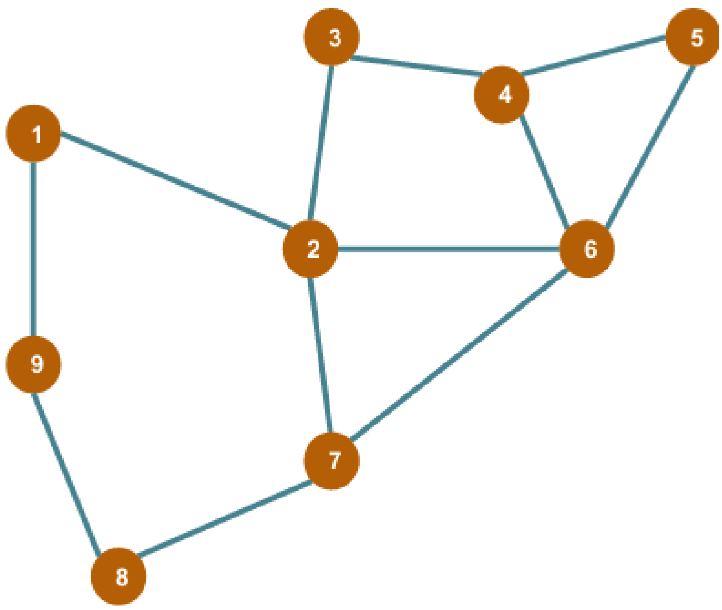
En el gráfico anterior, se definió el siguiente multigrafo: $G = \{\mathbf{(1,\ 2),\ (1,\ 2),\ (1,\ 2)}, (2,3), (3,4), (3,2), (4,3)...\}$

Pseudografo

Un Pseudografo es un tipo de **grafo no simple** el cual tiene la particularidad de que puede haber más de una arista entre el mismo par de vértices, y, además, los vértices pueden tener self-loops.
Good to Know: Puede haber más de un self-loop de un vértice.

Grado de un vértice (nodo)

El grado de un vértice es la cantidad de conexiones que tiene un nodo.
Definimos formalmente al grado de un vértice como $deg(v) = \#\{e \in E : v \in e\}$
Good to Know: deg refiere a degree.
Good to Know 2: **e** es una arista particular $e = (v, w)$.
Good to Know 3: En un grafo G, al grado mínimo lo notamos como $\delta(G)$ mientras que al grado máximo lo notamos como $\triangle(G)$



En este caso particular, tenemos que $\delta(G) = 2$ y $\triangle(G) = 4$, y, a modo de ejemplo, $deg(2) = 4$
Good to Know 4: Por cada par de vértices, tenemos una suma de grado 2. Ej.: $\{(1, 2)\}$ nos da $deg(1) + deg(2) = 2$

Grado de un vértice en un Multigrafo

Exactamente igual que en un Grafo, pero acá hay que aclarar que como podemos tener una relación entre dos vértices más de una vez tenemos que contar cada arista. Por ejemplo, si tenemos $\{(1,2),(1,2)\}$ entonces $deg(1) = 2$.

Grado de un vértice en un Pseudografo

Exactamente igual que en un Grafo, pero acá el self-loop cuenta 2 veces. Por ejemplo, si tenemos $\{(1,2),(1,2),(1,1),(1,1)\}$ entonces $deg(1) = 6$. Si en algún momento te maréas, el self-loop vale dos porque tenés el mismo número en el par.

Suma de los grados de los vértices

Dado un grafo G = (V, E) la **la suma de los grados** de sus vértices es igual a 2 veces el número de aristas. Es decir,

$$\sum_{v \in V} deg(v) = 2 \times \#\{e \in E = 2m\}$$

Para ver acerca de cómo demostrarlo, véase **anexo**
Corolario: En un grafo, siempre hay una cantidad par de vértices que tienen grado impar. Ej.: $\{(1,2),(2,3)\}, deg(1) = 1, deg(2) = 2, deg(3) = 1$

Grafo Dirigido (Digrafo)

Un Grafo Dirigido es un tipo de grafo (que puede ser simple o no simple) pero importa quién se relaciona con quién, es decir, el orden en que guardamos la relación. En dibujos, lo notamos con una flecha.

- $E = \{(A,B)_{\rightarrow}, (B,A)_{\leftarrow}\}$ es un grafo dirigido, esto quiere decir que $(A,B) \neq (B,A)$

Good to Know: A nivel estructura de datos, parece que E es igual en un Grafo Dirigido que un Grafo No Dirigido. No obstante, optamos, a nivel de notación incluir una fecha para indicar la relación entre ambos. A nivel código, deberías manejar alguna información extra para entender si es dirigido o no.

Grado de un vértice en un Grafo Dirigido

En un Grafo Dirigido, el grado de un vértice se calcula como $deg(v) = indeg(v) + outdeg(v)$
¿A qué nos referimos con indeg(v) y outdeg(v)? Como es un grafo dirigido, la relación entre los vértices importa. Por lo tanto, aquellas aristas entrantes al nodo v origen las agrupamos en indeg(v) mientras que aquellas aristas salientes desde el nodo v las agrupamos en outdeg(v).

- $E = \{(A,B)_{\rightarrow}, (B,A)_{\leftarrow}, (A,C)_{\rightarrow}\}$ es un grafo dirigido. Siendo indeg(A) = 1, outdeg(A) = 2 $\implies deg(A) = 3$

Cantidad de Aristas de un Grafo Dirigido Simple

Como acá importa la relación los vértices pues contamos todas las aristas pues $(A,B) \neq (B,A)$

$$0 \leq |E| \leq n(n-1)$$

Cantidad de Aristas de un Grafo Dirigido NO Simple

No hay límite, pueden ser infinitas. No hay fórmula concreta.

Grafo No Dirigido

Un Grafo No Dirigido es un grafo el cual no existe la dirección en una relación entre dos vértices.

- $E = \{(A,B),(B,A)\}$ es un grafo no dirigido, esto quiere decir que $(A,B) = (B,A)$

Good to Know: Los Grafos Dirigidos o Grafos no Dirigidos pueden ser Grafos Simples o no Simples.

Cantidad de Aristas de un Grafo NO Dirigido Simple

Como no importa la relación entre los vértices pues $(A,B) = (B,A)$, contamos solo una vez esa arista

$$0 \leq |E| \leq \frac{n(n-1)}{2} = \binom{n}{2}$$

Cantidad de Aristas de un Grafo NO Dirigido NO Simple

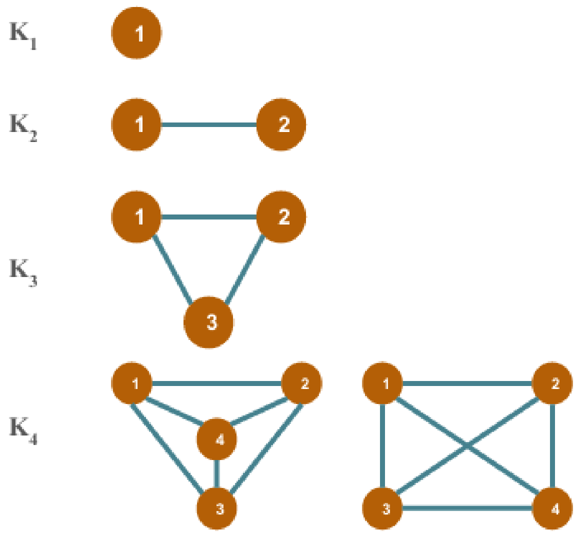
No hay límite, pueden ser infinitas. No hay fórmula concreta.

Grado de un vértice en un Grafo No Dirigido

Es la misma cuenta que hacemos en un grafo dirigido, pero sin distinguir en grupos a las aristas **(porque no tienen dirección)**.

Grafo Completo

Llamamos Grafo Completo a un Grafo que tiene todos sus vértices están conectados entre sí.
Formalmente, lo notamos como K_n donde **n** es la cantidad de vértices que tiene el grafo.



Grafo Complemento

Un Grafo Complemento (G^C) es el Grafo que tiene todas las aristas que no estaban en G.
Ej.: $G = \{(1,2),(3,2)\}$ entonces $G^c = \{(1,3),(2,3)\}$

Cantidad de aristas de un Grafo Complemento

$$m_{\overline{G}} = \frac{n(n-1)}{2} - m$$

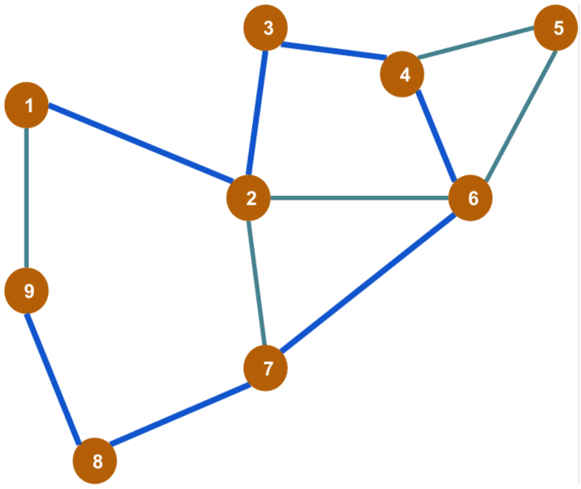
Creo que es necesario que sea un Grafo simple no dirigido

Recorrido

Un recorrido es una **sucesión de vértices** y **aristas** de un grafo, tal que e_i sea incidente a v_{i-1} y v_i para todo $i = 1...k : P = v_0e_1v_1e_2...e_kv_k$ (vértice, par de vértices, vértice, par de vértices, vértice...).
Ej.: $P = 1 \to (1,2) \to 2 \to (2,3) \to 3 \to (3,4) \to 4 \to (4,6) \to 6 \to (2,6) \to 2...$
Es decir, el nodo anterior (si hay) debe estar en la siguiente arista en alguna de las posiciones.

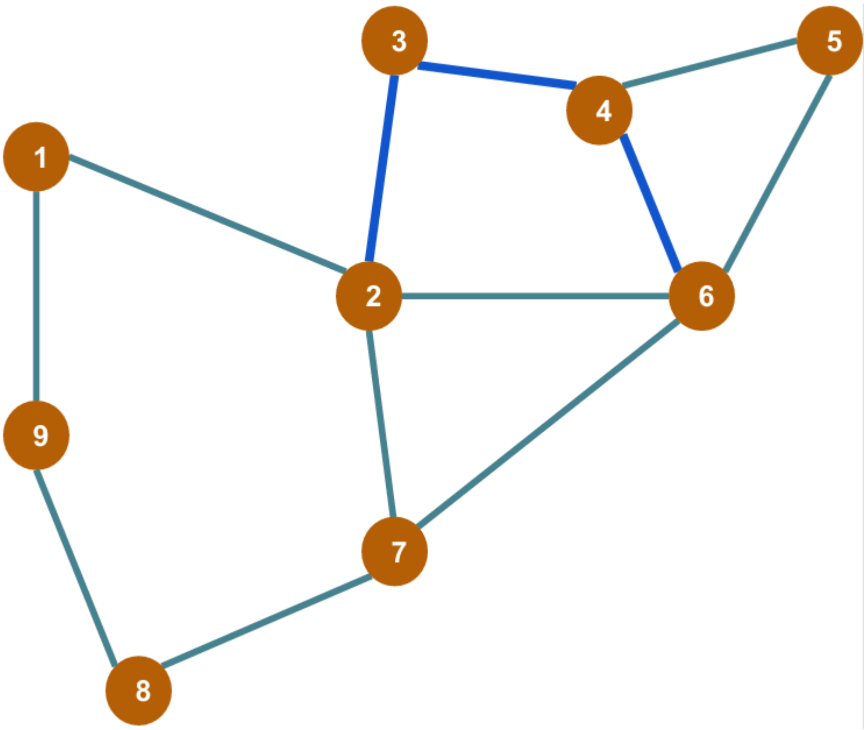
Camino

Un camino es un recorrido **que no pasa dos veces** por el mismo vértice.
Ej.: $P = 1 \to 2 \to 3 \to 4 \to 5 \to 6 \to 7 \to 8 \to 9$



Sección (subconjunto/parte de un recorrido)

Una sección es un tramo del recorrido P, se nota $P_{v_i} v_j$. Nótese que P es el recorrido, y los subíndices es aquello que acota al mismo.
Ej.: $P_{2,6} = 2 \to 3 \to 4 \to 6$. En este ejemplo, $v_i = 2$ y $v_j = 6$.

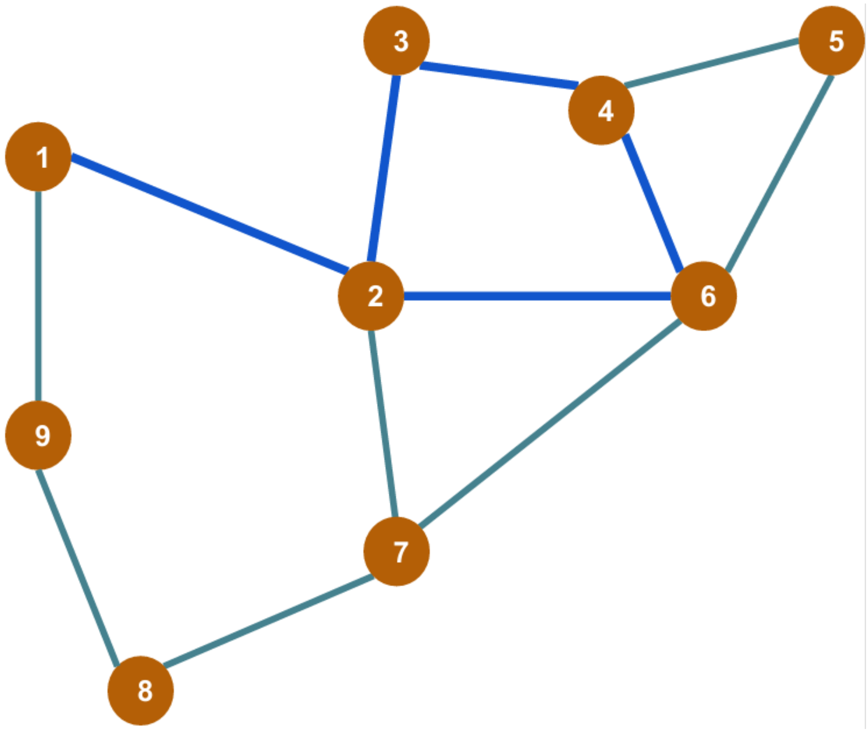


Good to Know

- Una sección es un subconjunto de un recorrido P.

Circuito

Un circuito es un recorrido que empieza y termina en el mismo vértice, se nota $P_{v_i}v_j$
Ej.: $P_{1,1} = 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 6 \rightarrow 2 \rightarrow 1$

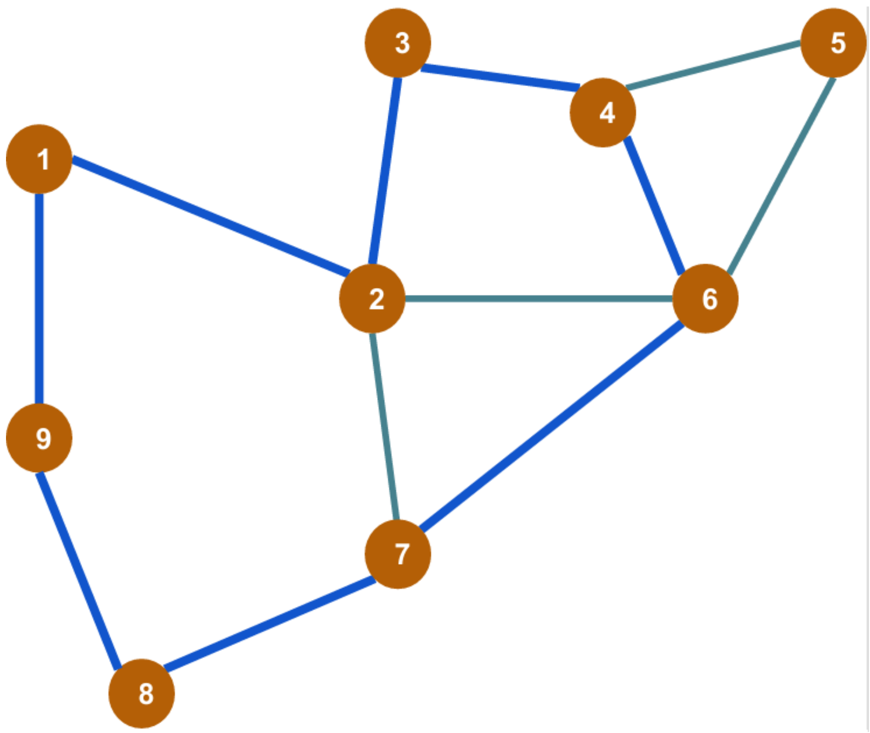


Good to Know:

- Un circuito es un tipo de recorrido.
- Un circuito es opuesto a un camino. Porque un camino no tiene repetidos.

Ciclo (o circuito simple)

Un ciclo o circuito simple es un circuito (de tres o más vértices) que no repite vértices, se nota $P_{v_i}v_j$. Ej.: $P_{1,1} = 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 6 \rightarrow 7 \rightarrow 8 \rightarrow 9 \rightarrow 1$

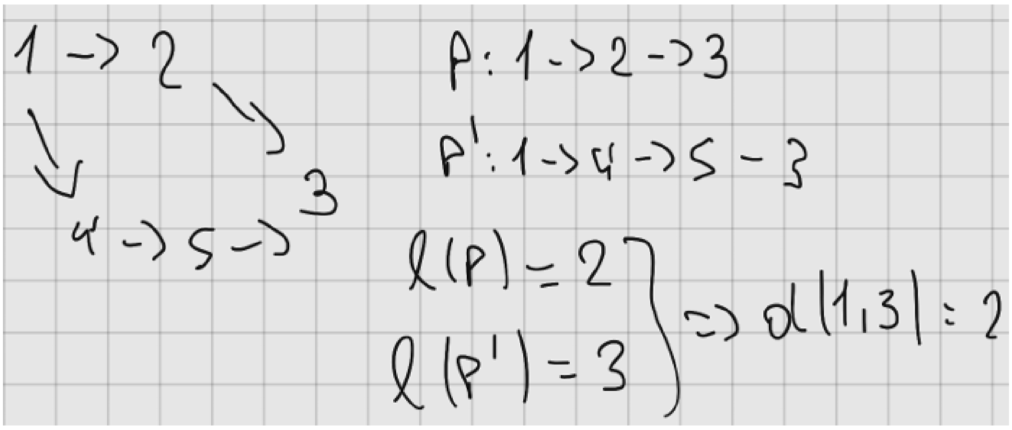


Longitud

La longitud de un recorrido P denotada l(P), es la **cantidad de aristas** que tiene.
Ej. : $P = 1 \to 2 \to 3 \to 4 \to 6 \to 2 \to 7$ donde $l(P) = 6$ **Good to Know**: Esta definición es muy importante, porque tranquilamente podría haber dos recorridos diferentes entre cada par de vértices.

Distancia entre dos vértices

La distancia entre dos vértices u y v se define como la **longitud del recorrido más corto** entre u y v. Se denota $d(u, v)$



Good to Know

- Precondiciones
  - Tener dos vectores: u y v que pertenezcan a un grafo G.
  - Tener uno o más recorridos, que incluyan los nodos de G.
- Postcondiciones
  - d(u,v) es el recorrido más corto desde u a v.
  - Si $u = v \iff d(u, v) = 0$
  - Si no existe un recorrido entre u y v entonces $d(u, v) = \infty$

Véase **anexo** para ver como demostrarlo por absurdo.

Camino en un Grafo Dirigido

Importa quién se relaciona con quien. Ej.: $E = \{(A, B)_{\to}, (B, C)_{\to}\}$

Camino en un Grafo no Dirigido

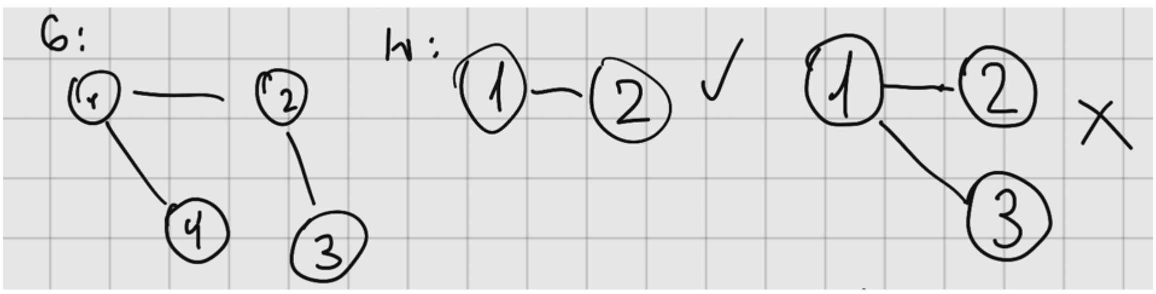
No importa quién se relaciona con quien. Ej.: $E = \{(A, B), (B, C)\}$. El camino podría ser en cualquier sentido.

Subgrafo

Sea $G = (V_G, E_G)$ y $H = (V_H, E_H)$

Subgrafo (base)

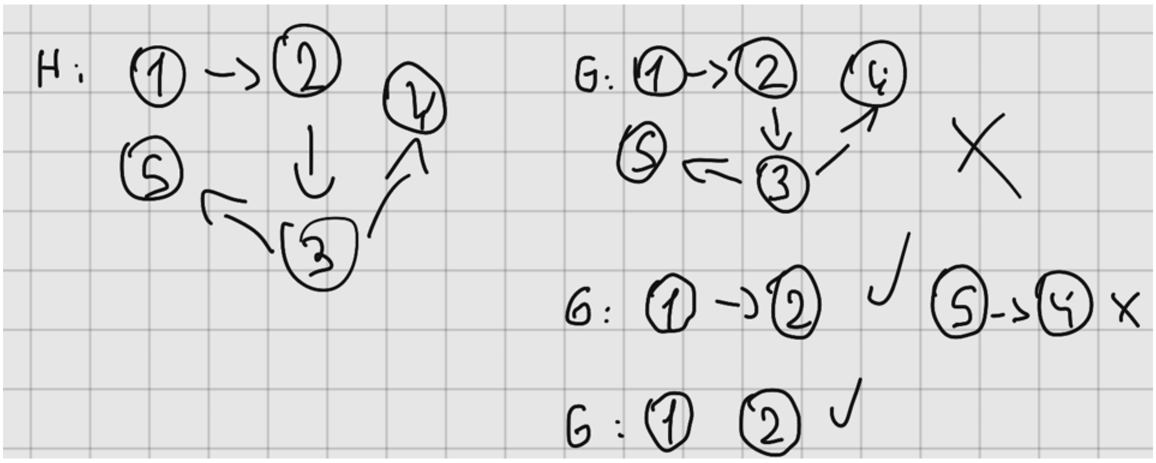
H es un subgrafo de G si todos los vértices y aristas de H están en G.
Esto quiere decir que con que no haya aristas (conexiones entre vértices) o vértices que no estén en G, H es un subgrafo de G.



A partir de acá, los demás subgrafos cumplen mínimamente estas condiciones

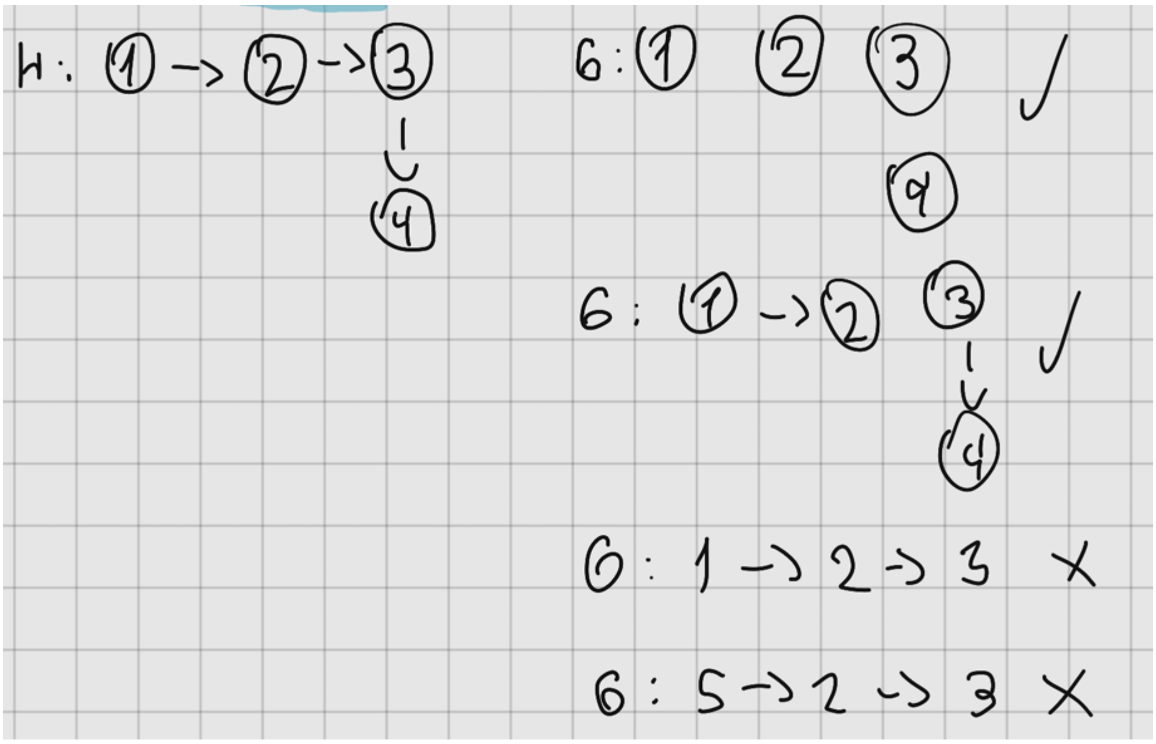
Subgrafo propio

H es un subgrafo propio de G si todos los vértices y aristas de H están en G siempre y cuando **no sea la totalidad**. Es decir, no debe suceder que sea exactamente $H = G$.



Subgrafo Generador

H es un subgrafo generador de G si H **tiene exactamente los mismos vértices** que G.



Un subgrafo generador es más fuerte que un subgrafo propio porque nos hacen usar todos los nodos pero conectarlos como queramos (siempre y cuando esas conexiones existan en el grafo original). Un subgrafo propio solo tiene una condición, no ser igual al grafo original.

Subgrafo Inducido

No sé por qué siento que este es trivial, o quizá lo estoy entendiendo mal.
H es un subgrafo inducido de G si todos los nodos de H, forman aristas que están en G y por consecuente, esas mismas aristas están en H.
Entiendo que es un subgrafo normal, solo que acá se obliga a que todos los nodos de H, tengan aristas que estén en G (subgrafo común), y por lo tanto están en H.
Lo que no entiendo es, ¿no es obvio que si hay aristas en H, deben estar sí o sí en G? ¿O acá lo que está obligando a que si están en G están si o sí en H?

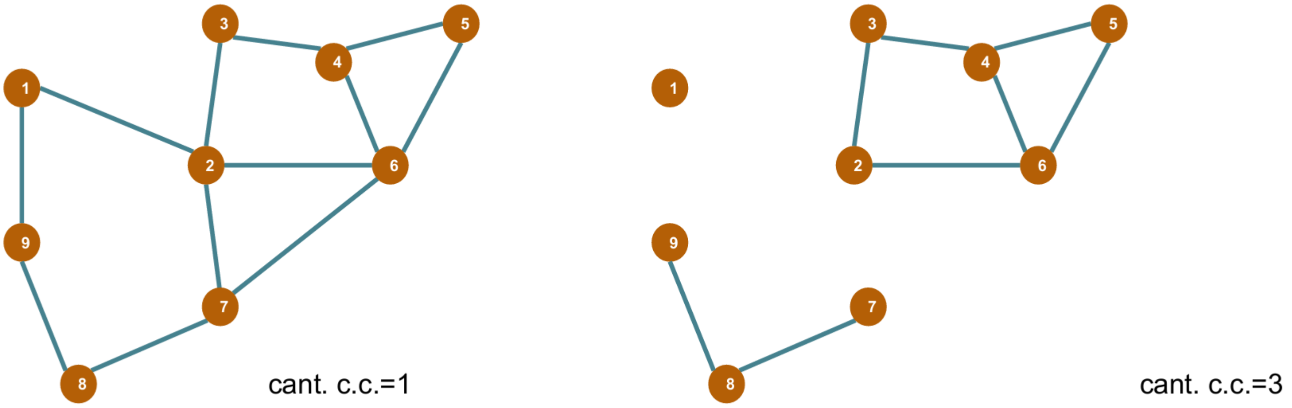
Grafo Conexo

Un Grafo Conexo es un grafo el cual existe un **camino** que conecta todos los vértices entre sí.

- $E = \{(A, B)_{\rightarrow}, (B, C)_{\rightarrow}, \}$ es un grafo conexo.
- $E = \{(A, B)_{\rightarrow}, (C, D)_{\rightarrow}\}$ no es un grafo conexo.

Componente Conexa

Una componente conexa es un subgrafo.



Normalmente se consiguen cuando desconectás alguna arista de un grafo conexo. Entonces, te quedan subgrafos. La cantidad de subgrafitos son los componentes conexos.

Grafo fuertemente Conexo

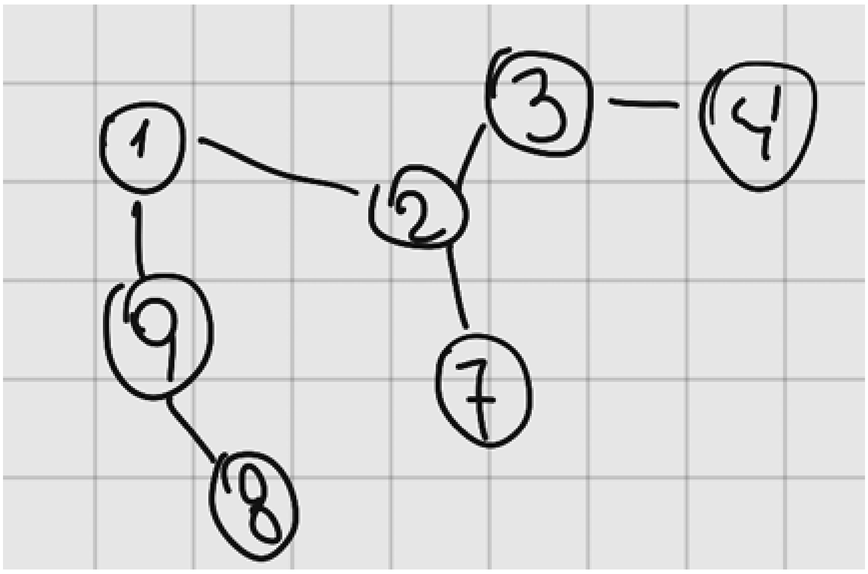
Un Grafo fuertemente conexo es un grafo dirigido y conexo. Todos los nodos deben tener una relación simétrica entre ellos.

- $E = \{(A, B)_{\rightarrow}, (B, C)_{\rightarrow}\}$ es un grafo conexo.
- $E = \{(A, B)_{\rightarrow}, (B, C)_{\rightarrow}, (C, B)_{\rightarrow}, (B, A)_{\rightarrow}, \}$ es un grafo fuertemente conexo.

Árboles

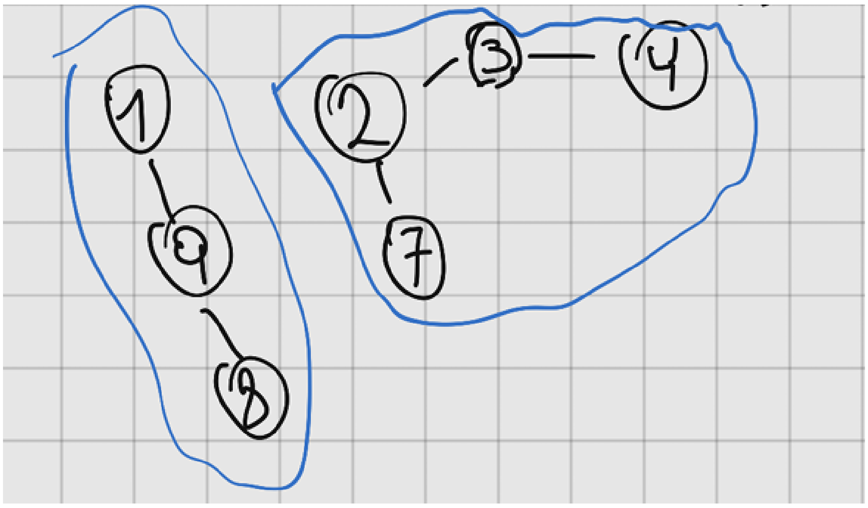
Base

Es un grafo conectado, sin ciclos (circuitos simples).



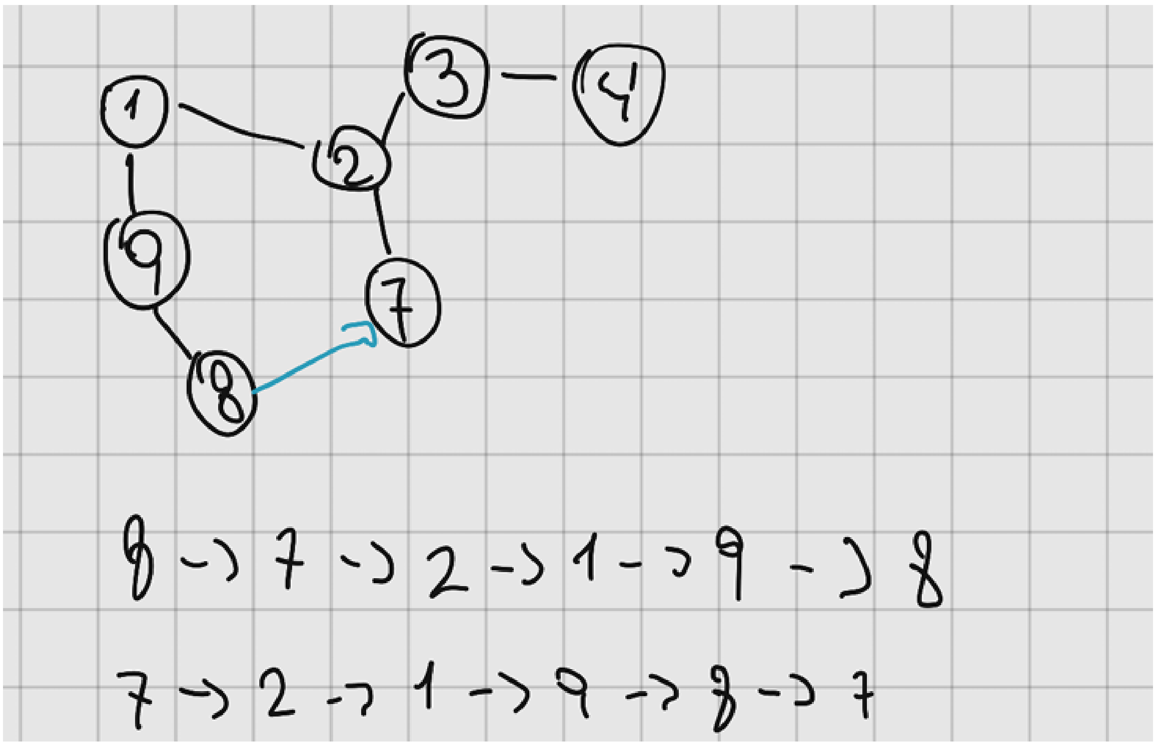
Quitar Arista

Si quitamos una arista cualquiera de un árbol, tenemos dos subgrafos conexos. Esto porque inicialmente un árbol no tiene ninguna componente suelta en el aire.



Agregar Arista

Si agrego una arista cualquiera se forma un ciclo (circuito simple).



Grafo T

Es un árbol que cumple con las siguientes propiedades

- Es conexo.
- No tiene ciclos.

Grafo Pesado

Definimos un Grafo Pesado G como $G = (V, E, W)$ donde W es una función que recibe dos vértices y devuelve el peso. Ahora, en cada arista almacenamos ambos vértices y el resultado del peso.

$$E_i = (v_i, v_j, w(v_i, v_j))$$

Grafo Acíclico Dirigido (DAG)

Un Grafo Acíclico Dirigido es un grafo dirigido que no tiene ciclos. Es decir, no existe ningún vértice que es capaz de empezar y terminar un recorrido.
Ej.: Una Single Linked List es un caso particular de un DAG.

Grafos Bipartito

Dado un grafo $G = (V, E)$, G es bipartito si

- Existe V_1 y V_2 tal que $V_1 \cup V_2 = V$
- $V_1 \cap V_2 = \emptyset$
- $e = (u, v) \in E, u \in V_1, v \in V_2$. Esto significa que el vértice de V_1 se relaciona con el de V_2 o en el otro sentido.
- La única restricción es que no podés conectar un vértice de V_1 con otro de V_1 ni tampoco uno de V_2 con otro de V_2

Grafo Bipartito Completo

Lo mismo que un Grafo Bipartito pero acá todos los vértices de V_1 están conectados con los de V_2 y viceversa.
Básicamente, el producto cartesiano.

Cosas a releer / preguntar

- Un grafo es bipartito, $V = (V_1, V_2) \iff$ no tiene ciclos de longitud impar.
- Un grafo es bipartito $\iff$ todas sus c.c. lo son
- Un grafo no tiene ciclos impares $\iff$ cada una de sus c.c. no tienen ciclos impares.

Isomorfismo de Grafos

Dados $G = (V, E)$ y $G' = (V', E')$. G y G' son isomorfos si existe una función biyectiva $f : V \to V$ tal que $\forall u, v \in V$

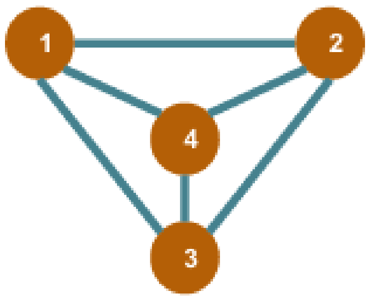
$$(u, v) \in E \iff (f(u), f(v)) \in E'$$

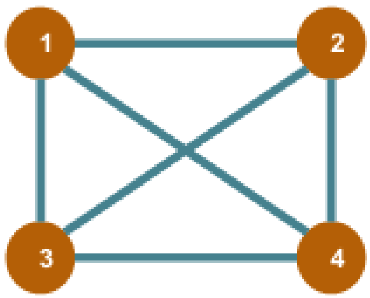
Es decir, **si parto desde E y agarro el par que está en una arista, puedo poner cada componente en una función f y me lo transforma en una arista que está en E'**.
Lo mismo con el caso opuesto.

Propiedades

Si dos grafos $G = (V, E)$ y $G' = (V', E')$ son isomorfos entonces

- Tienen el mismo número de vértices (sí, porque en la biyectividad se cumple la inyectividad y la sobreyectividad).
- Tienen el mismo número de aristas (sí, porque cada una debería traducirse a una única específica, de lo contrario no sería inyectiva).
- Para todo k, $1 \leq k \leq n - 1$, tienen el mismo número de vértices de grado k (la misma distribución de grado).
- Tienen la misma cantidad de componentes conexas.
- Para todo k, $1 \leq k \leq n - 1$, tienen el mismo número de caminos simples de longitud k.

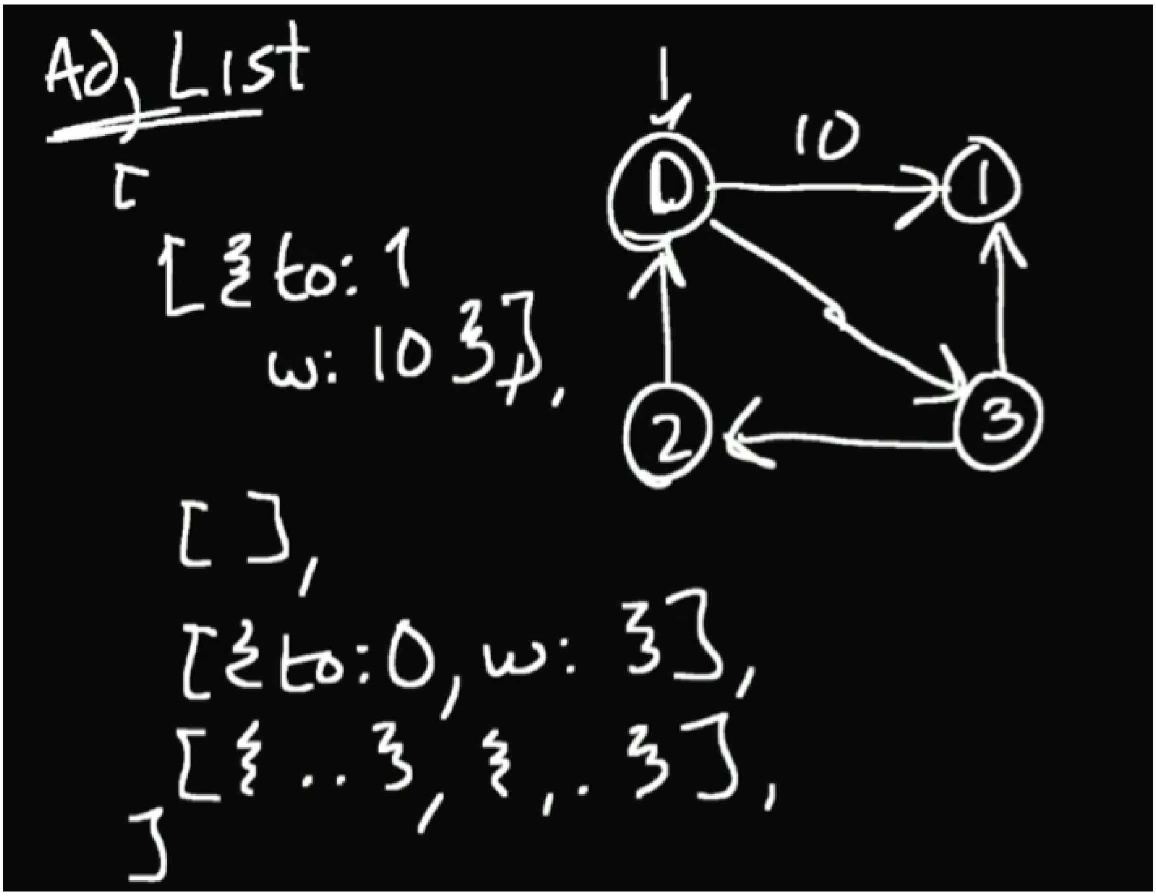




Representación de grafos

Lista de Adyacencia

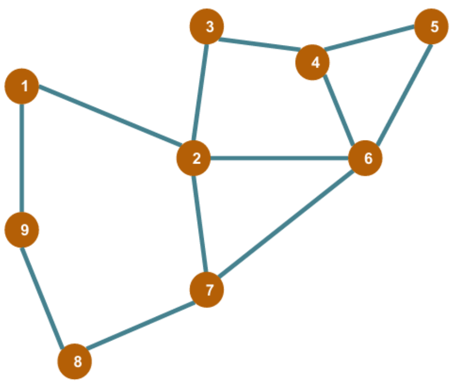
Es la más común y la más eficiente.
Es una lista de listas, donde cada lista corresponde al vértice y las aristas con las que se conecta se representa con una lista de diccionarios.



Matriz de Adyacencia

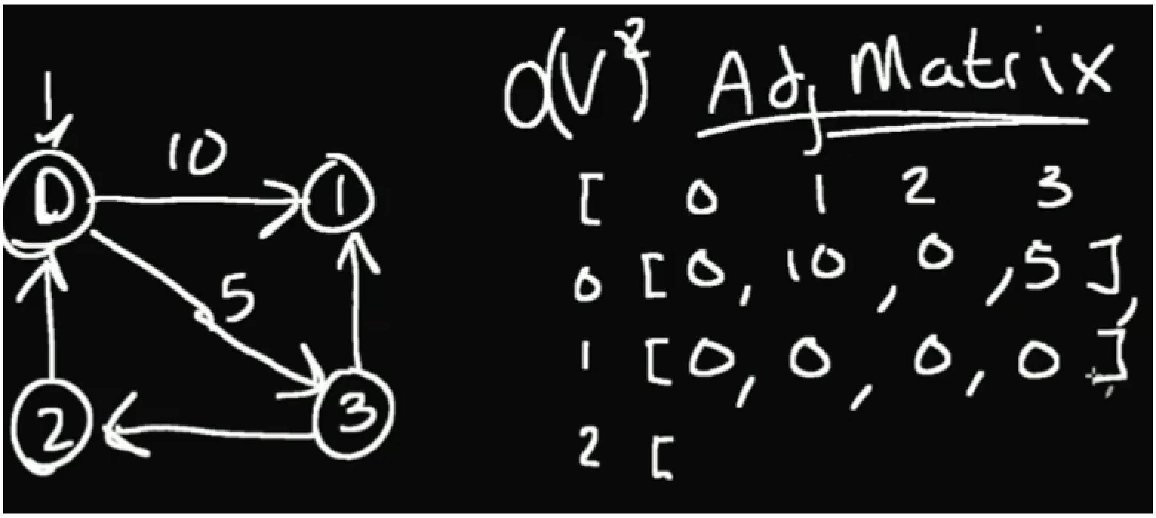
Cada número está representado por el índice en donde está. Vale 1 **(o el valor de peso de la arista)** si existe la arista y 0 si no.
La cantidad de filas y columnas es igual a la cantidad de nodos que hay, es por eso que espacialmente cuesta $O(n^2)$

Matriz de adyacencia (a_{ij})



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

Ej.: El nodo 3 (fila 3), está relacionado con todos los nodos que están en la columna 3. En el dibujo, el nodo 3 está relacionado con el 2 y con el 4. Por eso, $a_{3,2} = 1$ y $a_{3,4}$.



Good to Know: En el dibujo vemos que empieza todo desde el índice 1 pero en realidad empezamos desde el 0. **Lo que todavía no entiendo es, qué pasaría si los nodos guardan información compleja, como elijo cual va en el índice 0, cual en el 1, etc.**
Esta forma de trabajar con grafos es **súper ineficiente**.

Anexo

Demostraciones de Grafos

Normalmente se hace inducción sobre las aristas, aunque en algunos casos, optamos sobre los vértices.
Lo más común es tener un caso base del tipo $n = 1, m = 0$ o $n = 2, m = 1$. **Good to Know**: Usar a favor las propiedades.

Distancia entre dos vértices

Si P es el recorrido entre u y v tiene longitud d(u,v) $\implies$ P es un camino.
Recordemos algunas definiciones

- Un recorrido es una sucesión de vértices y aristas, que no tiene ninguna limitación.
- d(u, v) es el recorrido más corto entre u y v. Por lo que, podría haber otros.
- P es un camino: Un camino es un subconjunto de un recorrido donde no puede haber vértices repetidos.

Probando por el absurdo:

$$\neg Q \implies R$$

En este caso

- $\neg Q =$ P no es un camino
- $R =$ P es el recorrido entre u y v y tiene longitud d(u,v)

La idea es llegar a algo falso.

$$V \implies F = F$$

Suponemos que P NO es un camino $\neg Q = V$, entonces hay, al menos un vértice en P que se repite.
Recorrido P

$$u \to z \to z \to v$$

¿Podemos encontrar algún camino que sea más corto que P? (esto haría falso el consecuente).
Defino un camino T

$$u \to z \to v$$

Luego, la distancia mínima entre ambos recorridos

$$min(dP(u, v), dT(u, v)) = min(3, 2) = 2 = dT(u, v)$$

Finalmente, llegamos a un absurdo, porque la distancia mínima la tenemos con el recorrido T y no con P.
Si hubiesemos llegado a que con $\neg Q$, P es el camino más corto, hicimos mal la demostración.

Suma de los grados de los vértices