

Estructura de Datos

Semana 12



Logro de la sesión

Al finalizar la sesión, el estudiante:

 Comprenderá el concepto de grafo y sus principales propiedades. Implementará las diferentes representaciones de un grafo así como los diferentes algoritmos de recorridos, finalmente resolverá problemas utilizando grafos.

Estructuras de datos no lineales

Grafos



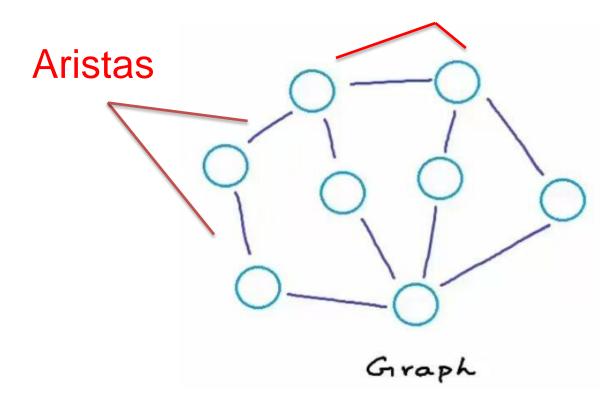
Agenda

- Árboles 2-3
- Árboles B
- Código de Huffman

Introducción

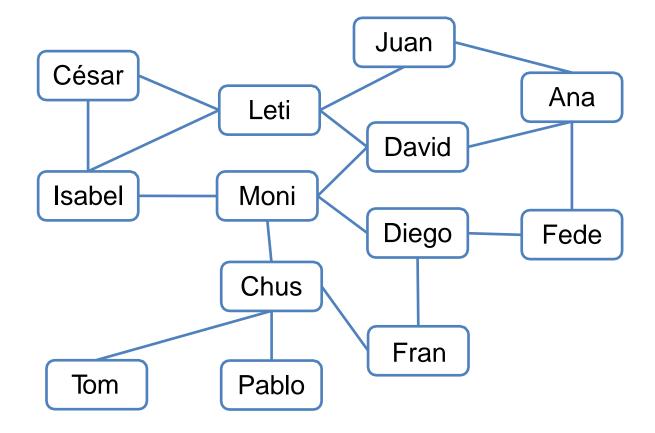
Estructuras no lineales: grafos

Nodos o vértices



Permiten representar cualquier tipo de conexión

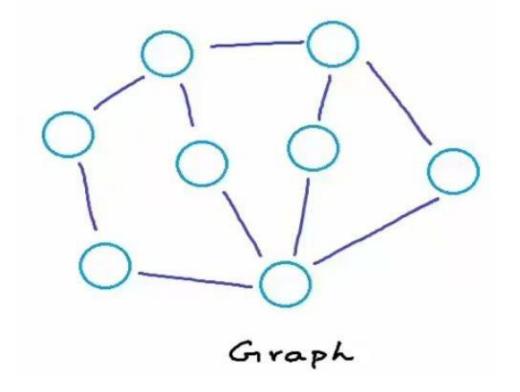
Introducción

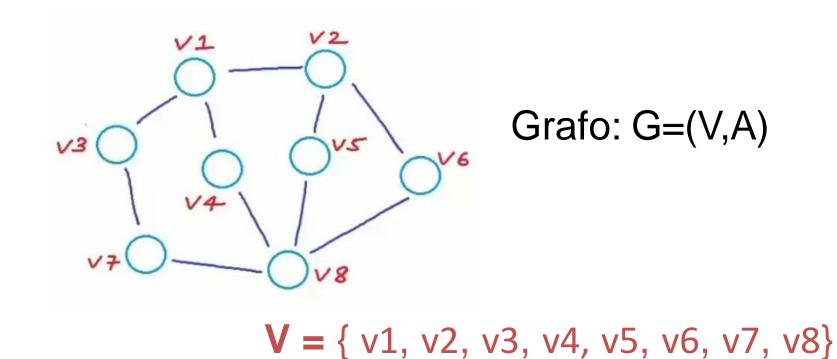


Una red social se puede representar como un grafo no dirigido y no ponderado. Los usuarios son los vértices y sus relaciones de amistad las aristas.

Grafo: G=(V,A)

Un grafo G es un par (V,A), donde V es un conjunto de vértices (nodos) y A un conjunto de aristas





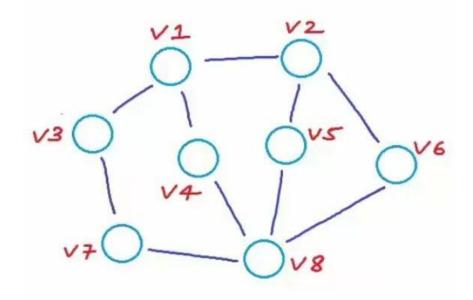
¿Cómo puedo representar una arista?

Tipos de Aristas

directed undirected
$$(u,v) \\ (u,v)!=(v,u) \text{ if } u!=v \\ \{u,v\}=\{v,u\}$$

Grafo: G=(V,A)

```
V = { v1, v2, v3, v4, v5, v6, v7, v8}
A= { {v1, v2}, {v1, v3}, {v1, v4}, {v2,v5}, {v2,v6}, {v3,v7}, {v4,v8}, {v5,v8}, {v6,v8}, {v7,v8}}
```

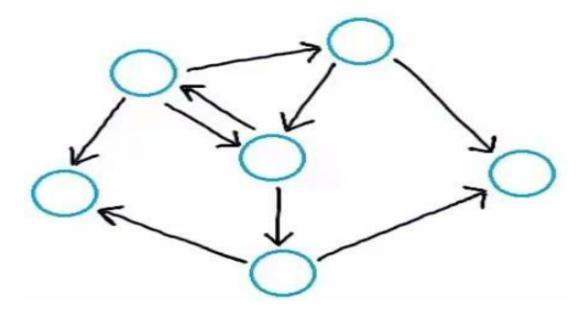


|V|=número de vértices

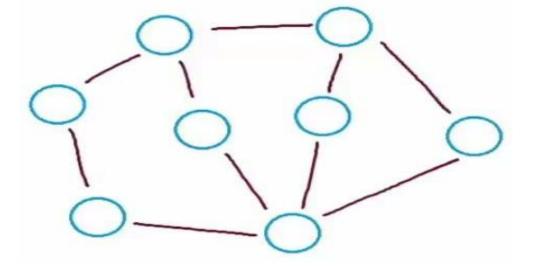
|A|=número de aristas

$$|A| = 10$$

Grafo dirigidos



Grafo no dirigido



Tipos de Grafos

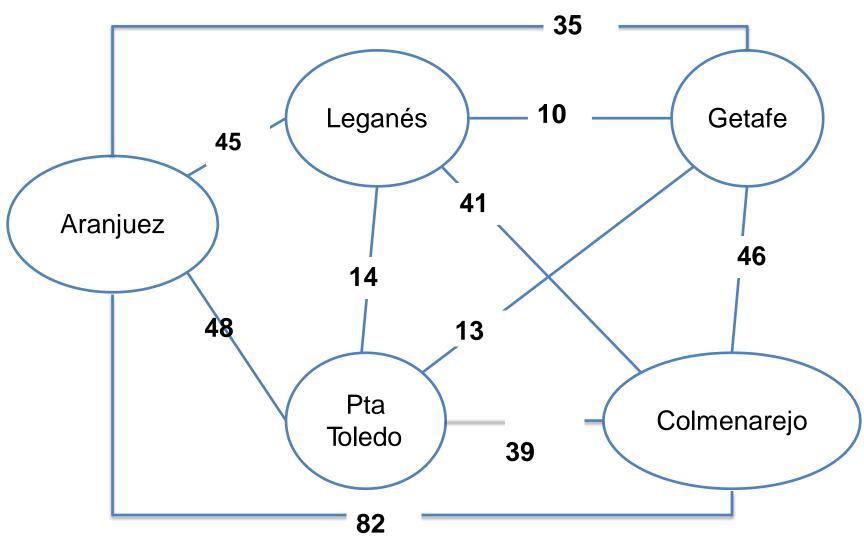
Grafos no dirigidos.

- Las aristas no tiene dirección, es decir, (u,v)=(v,u). La arista se puede recorrer en ambos sentidos.
- Nos permiten representar relaciones simétricas y de colaboración.

Grafos dirigidos.

- Cada arista (u,v) tiene una única dirección, siendo u el vértice origen y v el vértice final. (u,v) ≠(v,u)
- Nos permiten representar relaciones asimétricas y jerárquicas.

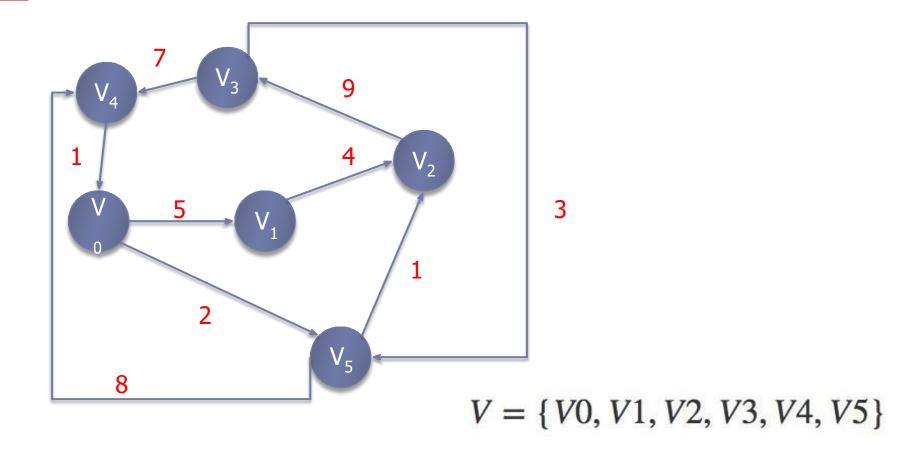
Grafo ponderado (weighted)



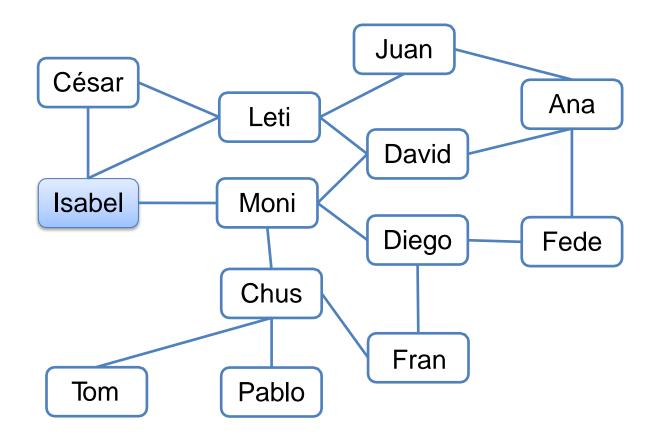
Conceptos Básicos

- ▶ Grafo Gdonde G=(V,A).
 - V es un conjunto de vértices (nodos)
 - ► A es un conjunto de aristas (arcos).
 - Una arista es una conexión entre dos vértices.
 - ► Cada arista puede ser representada como una tupla (v,w)donde $w,v \in V$
 - Además, cada artista puede tener un peso asociado (grafo ponderado). En este caso, la arista quedaría representada por una terna (v,w,p) donde p es el peso asociado a la arista entre v y w.

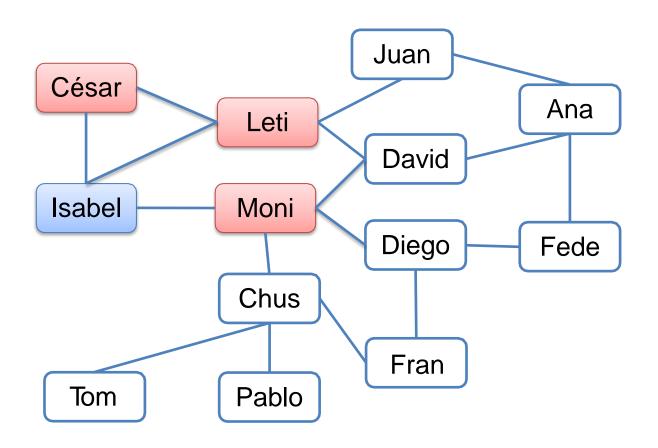
Grafo Ponderado

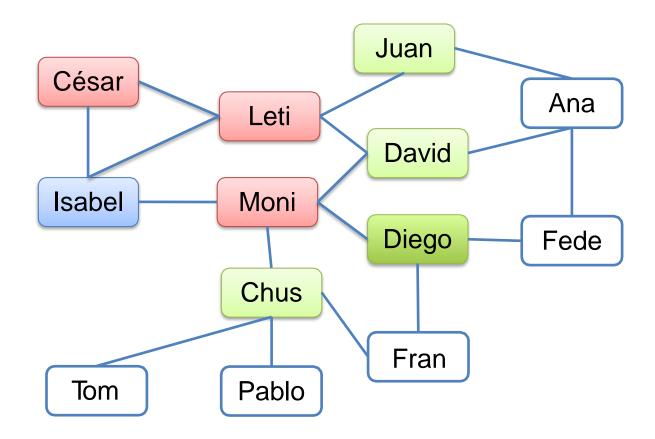


$$A = \left\{ \begin{array}{l} (v0, v1, 5), (v1, v2, 4), (v2, v3, 9), (v3, v4, 7), (v4, v0, 1), \\ (v0, v5, 2), (v5, v4, 8), (v3, v5, 3), (v5, v2, 1) \end{array} \right\}$$

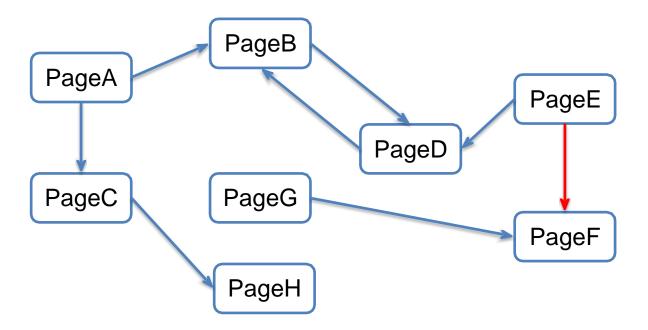


¿Cómo sugerir nuevos amigos a Isabel?

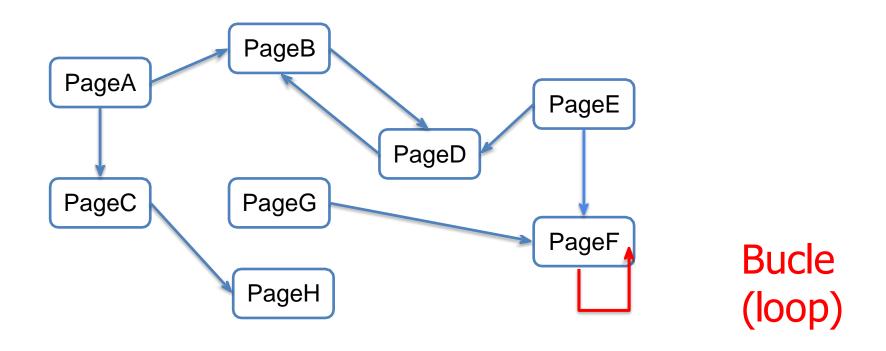




Encontrar todos los nodos para los que exista un camino de longitud 2

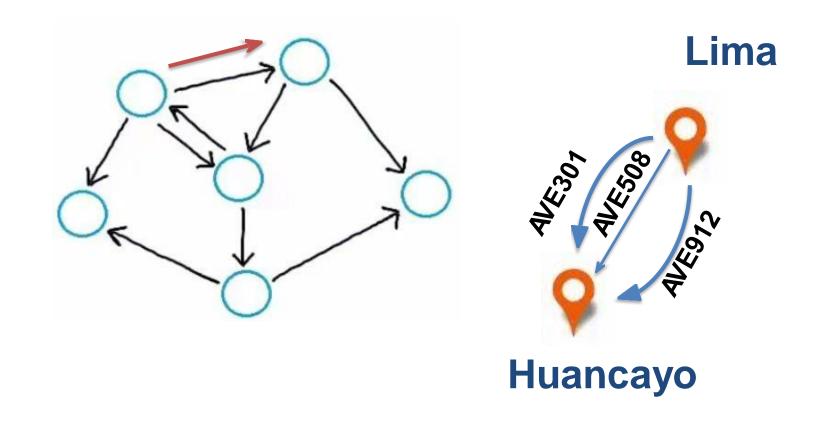


La Web se puede representar como un grafo dirigido. Los vértices son las páginas web y las conexiones entre estas son las aristas del grafo.



Una página puede contener un enlace a ella misma. Ese tipo de aristas son conocidas como bucles (loop) y son aristas que conectan un vértice consigo mismo.

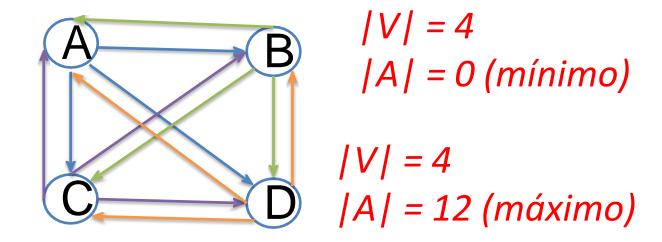
Aristas múltiples (aristas paralelas)



 Los bucles y las aristas paralelas tienden a hacer más complejos los algoritmos de grafos.

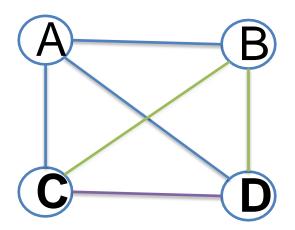
 Un grafo simple es un grafo que no tiene bucles ni aristas paralelas.

• ¿Cuál es el número mínimo y máximo de aristas en un grafo simple dirigido?



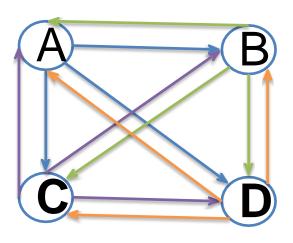
Si /V/ = n, cada vértice podría tener un máximo de n-1 aristas. Por lo tanto, 0<= /A/<=n(n-1)

 ¿Cuál es el número máximo de aristas en un grafo simple no dirigido?

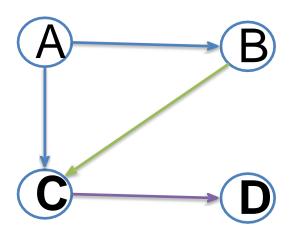


Si /V/=n, cada vértice podría tener n-1 aristas. 0 <= /A/<=n(n-1)/2

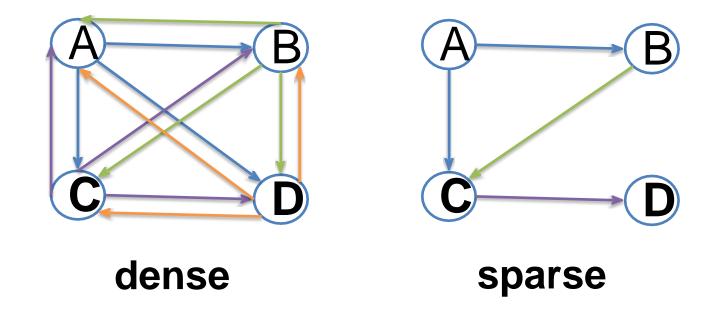
 Un grafo es denso si el número de sus aristas es cercano a su número máximo posible (n(n-1) o n(n-1)/2) (≈ n²)



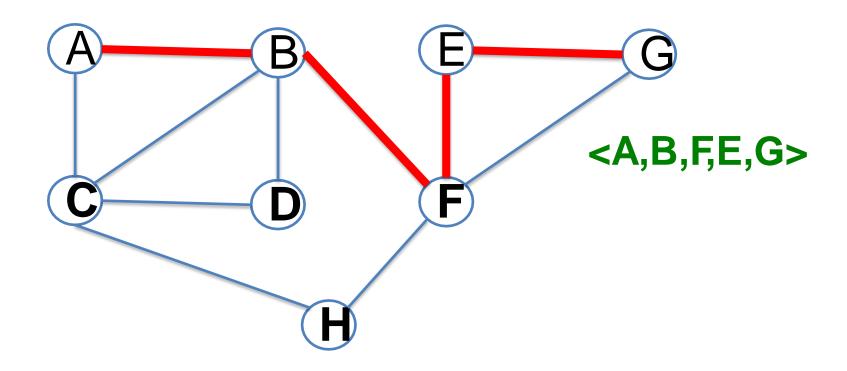
 Un grafo es disperso si el número de sus aristas es cercano a el número de sus vértices (≈n)



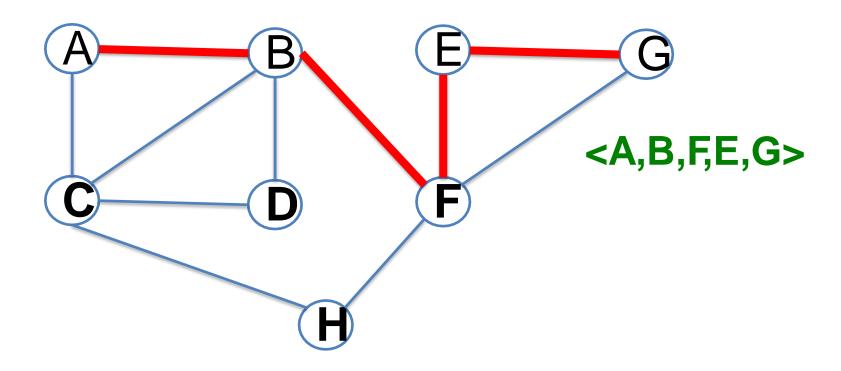
• Conocer si un grafo es denso o disperso nos ayudará a elegir la implementación más apropiada.



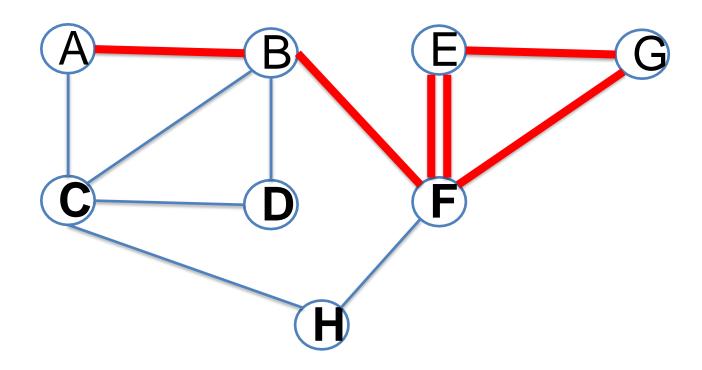
• Un camino es una secuencia de vértices tal que exista una arista entre cada vértice y el siguiente.



• Un camino simple es aquel que no repite vértices en su recorrido.



Este camino no es simple porque hay dos vértices repetidos



TAD Grafo

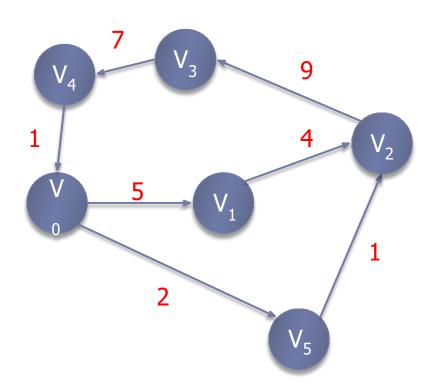
```
public interface IGraph {
    //devuelve el número de vértices
    public int sizeVertices();
    //devuelve el número de aristas
    public int sizeEdges();
    //muestra los vertices y sus aristas
    public void show();
    //devuelve el grado del vértice i
    public int getDegree(int i);
    //devuelve el grado de entrada del vértice i
    public int getInDegree(int i);
    //devuelve el grado de salida del vértice i
    public int getOutDegree(int i);
```

TAD Grafo (cont.)

```
//crea un nuevo vértice
public void addVertex();
//añade una arista entre i y j
public void addEdge(int i, int j);
//añade una arista entre i y j con peso w
public void addEdge(int i, int j, float w);
//borra la arista entre i y j
public void removeEdge(int i, int j);
//comprueba si existe una arista entre i y j
public boolean isEdge(int i, int j);
//devuelve el peso asociado a la arista (i,j).
public Float getWeightEdge(int i, int j);
//devuelve un array con los vértices adyacentes a i
public int[] getAdjacents(int i);
```

Implementación basada en matriz

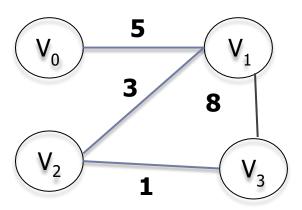
▶ La matriz de adyacencias



	V_0	V_1	V_2	V_3	V_4	V_5
V_0		5				2
V_1			4			
V_2				9		
V_3					7	
V_4	1					
V ₅			1			

Implementación – Matriz de adyacencias

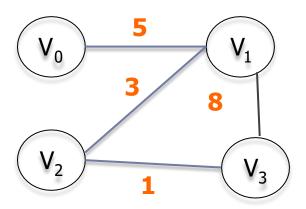
- Un grafo puede ser representado como una matriz cuadrada nxn, siendo n el número de vértices del grafo.
- Cada vértice v es representado por un entero (índice de v), en el rango {0,1,..,n-1} siendo n el número de vértices



- V₀ -> indice 0
- V₁ -> índice 1
- V_2 -> indice 2
- V_3^- -> indice 3

Implementación – Matriz de adyacencias

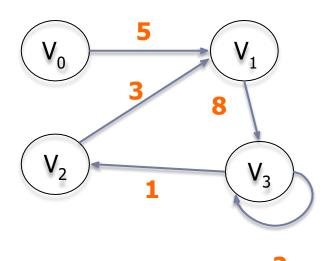
La matriz se puede implementar como un array bidimensional n x n M, tal que el elemento M[i,j] guarda información sobre la arista (v,w), si existe, donde v es el vértice con índice i y w es el vértice con índice j.



	0	1	2	3
0		5		
1	5		3	8
2		3		1
3		8	1	

Si el grafo no es dirigido la matriz es simétrica

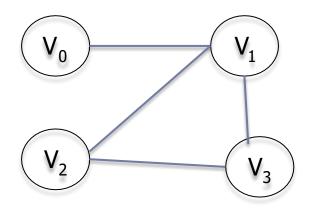
Implementación – Matriz de adyacencias



Si el grafo es dirigido la matriz NO es simétrica

	0	1	2	3
0		5		
1				8
2		3		
3			1	2

 Si es un grafo no etiquetado, el grafo se podría representar con una matriz de booleanos,

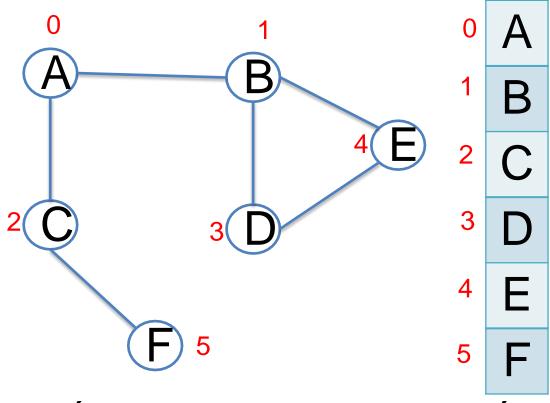


	0	1	2	3
0	false	true	false	false
1	true	false	true	true
2	false	true	false	true
3	false	true	true	false

Si el grafo no es dirigido la matriz es simétrica

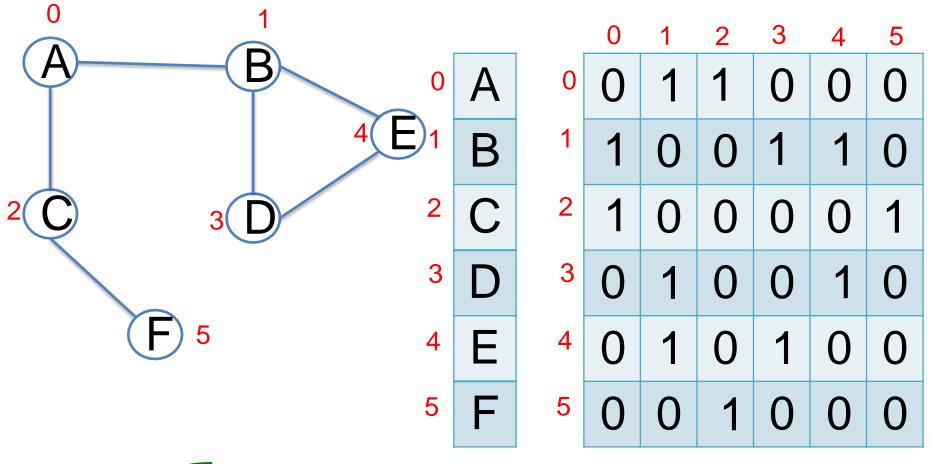
Matriz de Adyacencia

Lista de vértices



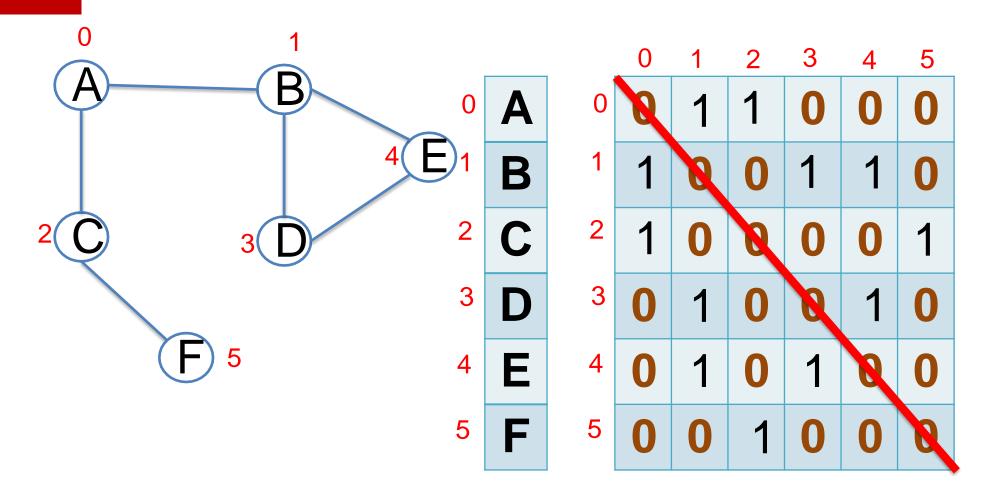
Cada vértice es representado por un índice. Podemos usar una lista de Python (array) para almacenar los vértices.

Matriz de Adyacencia (grafo no dirigido)



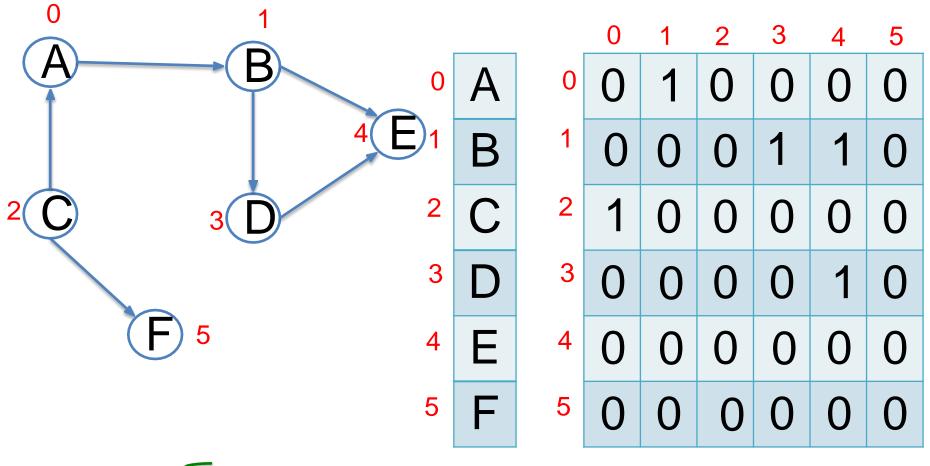
$$Mij = \begin{cases} 1, & \text{if } \{i, j\} \text{ si es una arista} \\ 0, & \text{eoc} \end{cases}$$

Matriz de Adyacencia (grafo no dirigido)



En los grafos no dirigidos, la matriz de adyacencia va a ser simétrica: Mij = Mji

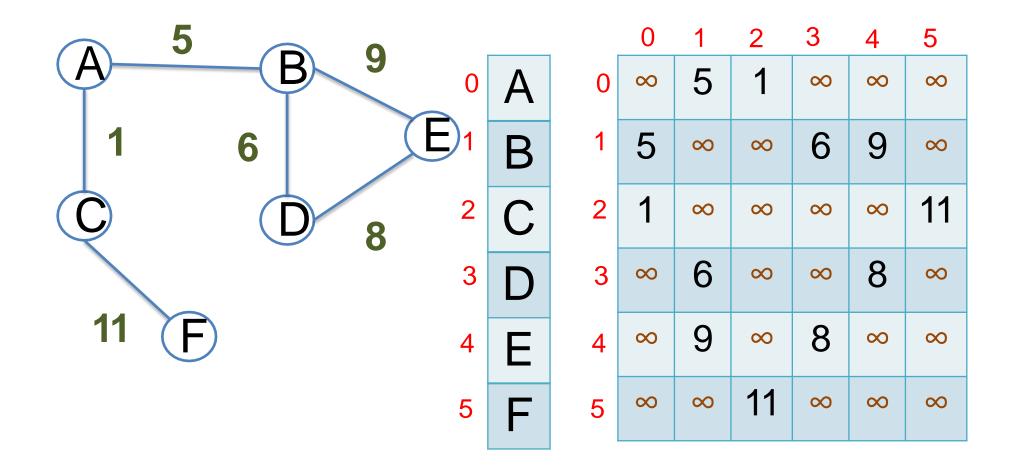
Matriz de Adyacencia (grafo dirigido)



Mij =
$$\begin{cases} 1, & \text{if (i, j) si es una arista} \\ 0, & \text{eoc} \end{cases}$$

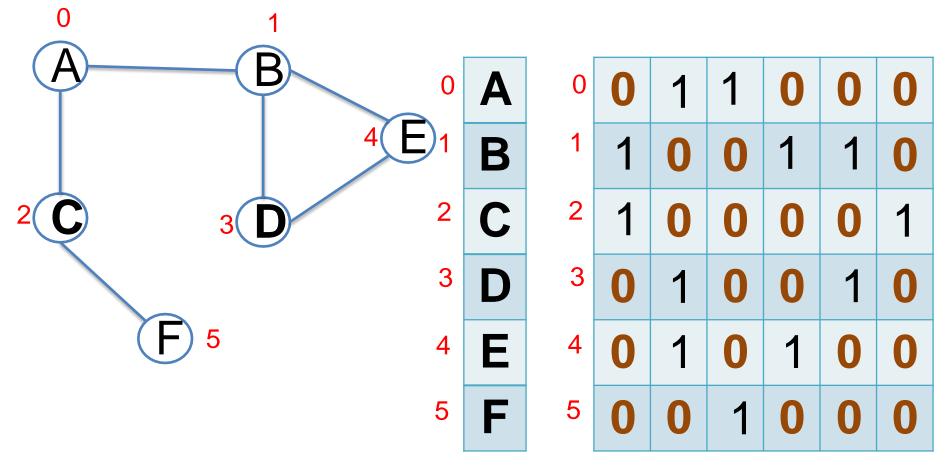
En este caso, la matriz no es simétrica

Matriz de Adyacencia (grafo no dirigido)



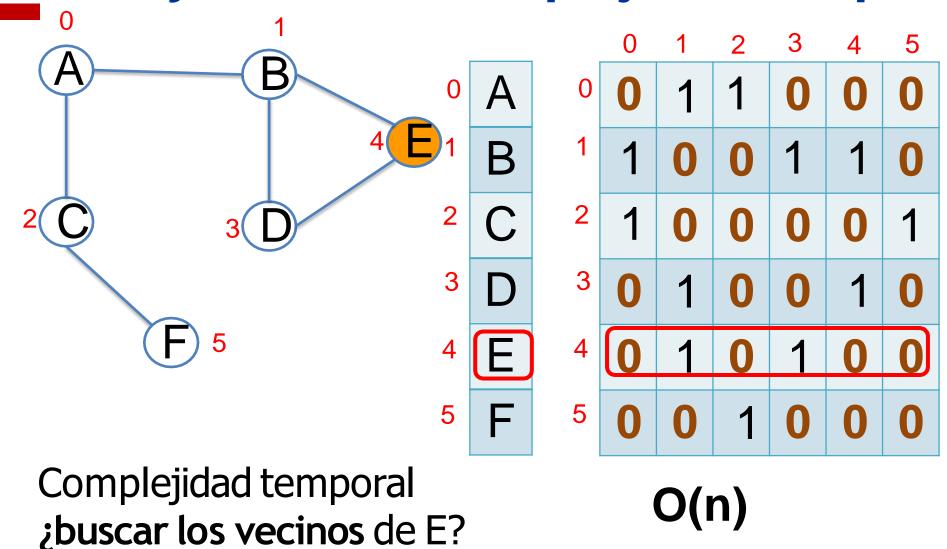
¿Cómo representar grafos ponderados?

Matriz de Adyacencia - Complejidad Espacial

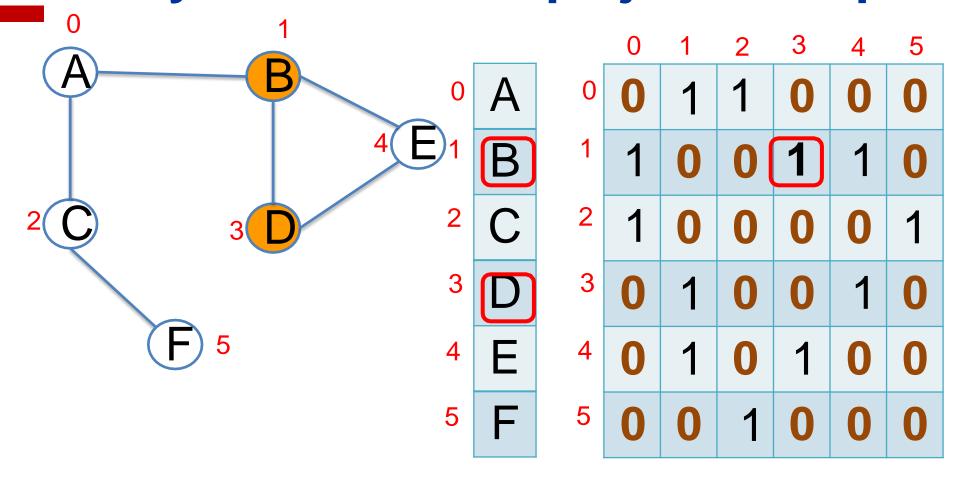


Complejidad espacial If |V| = n, $O(n^2)$

Matriz de Adyacencia - Complejidad Temporal



Matriz de Adyacencia - Complejidad Temporal



ison vecinos?

$$O(n) + O(1)$$

- ► Vamos a ver una implementación para un grafo no ponderado.
- La matriz puede ser almacenada en un array bidimensional de booleanos (true indicará que existe arista y false que no existe).
- La creación de nuevos vértices podría implicar la necesidad de modificar el tamaño asignado a la matriz.
- Para evitarlo, vamos a definir un atributo que almacene el número máximo de vértices (en ningún caso, se permitirá añadir un nuevo vértice cuando ese umbral se haya alcanzando) y otro atributo que almacene el número actual de vértices.

```
public class GraphMA implements IGraph {
    boolean matrix[][];
    //maximum number of vertices
    int maxVertices;
    //current number of vertices
    int numVertices;
    //true if the graph is directed, false eoc
    boolean directed;
```

```
public GraphMA (int n, int max, boolean d) {
    //We checks if the values are right for the graph
    if (max <= 0)
        throw new IllegalArgumentException("Negative maximum number of vertices!!!");
    if (n \le 0)
        throw new IllegalArgumentException("Negative number of vertices!!!.");
    if (n>max)
        throw new IllegalArgumentException("number of vertices can never be greater than the maximum.");
    maxVertices=max;
    numVertices=n;
    matrix=new Float[maxVertices][maxVertices];
    directed=d;
```

- Primero deberemos comprobar que todos los argumentos del constructor reciben valores apropiados: tanto el número máximo como el número de vértices debe ser siempre un número positivo.
- Además, el número de vértices nunca deberá sobrepasar el número máximo de vértices.
- El constructor crea el array bidimensional. Por defecto, todos las posiciones son inicializadas a false.

```
public void addVertex() {
    if (numVertices==maxVertices) {
        System.out.println("Cannot add new vertices!!!");
        return;
    }
    numVertices++;
}
```

- Lo primero que tenemos que hacer es comprobar que el nuevo número total de vértices no va a sobrepasar el número máximo permitido.
- Por último, sólo tendremos que incrementar en uno el número actual de vértices.
- No hace falta inicializar la matriz para el nuevo vértice, porque por defecto todas sus posiciones en la matriz son false.

```
//check if i is a right vertex
private boolean checkVertex(int i) {
   if (i>=0 && i<numVertices) return true;
   else return false;
}</pre>
```

- Vamos a usar un método auxiliar para comprobar si un índice representa o no un vértice en el grafo.
- Para que sea un vértice del grafo siempre deberá ser positivo y menor que numVertices, porque los vértices del grafo toman valores en el rango [0, numVertices-1].

```
public void addEdge(int i, int j) {
    if (!checkVertex(i))
        throw new IllegalArgumentException("Nonexistent vertex " + i);
    if (!checkVertex(j))
        throw new IllegalArgumentException("Nonexistent vertex " + j);
    matrix[i][j]=true;
    if (!directed) matrix[j][i]=true;
}
```

Una vez comprobado que ambos índices son correctos, simplemente lo que tenemos que hacer es actualizar la posición matrix[i,j] a true. Si no es dirigido, también tendremos que poner su posición simétrica

```
@Override
public boolean isEdge(int i, int j) {
    //checks if the indexes are right
    if (!checkVertex(i))
        throw new IllegalArgumentException("Nonexistent vertex " + i);
    if (!checkVertex(j))
        throw new IllegalArgumentException("Nonexistent vertex " + j);
    return matrix[i][j];
}
```

- En primer lugar, tenemos que comprobar que los índices i y j son correctos.
- El par (i,j) es un arista si matrix[i,j] guarda true.

```
@Override
public void removeEdge(int i, int j) {
    //checks if the indexes are right
    if (!checkVertex(i))
        throw new IllegalArgumentException("Nonexistent vertex " + i);
    if (!checkVertex(j))
        throw new IllegalArgumentException("Nonexistent vertex " + j);
    matrix[i][j]=false;
    if (!directed) matrix[j][i]=false;
}
```

- Primero tenemos que comprobar que son índices válidos
- Una vez comprobado que son índices válidos, basta con modificar el valor del array en esa posición (i,j) a false.
- Si no es dirigido, también tendremos que hacerlo en su elemento simétrico (j,i)

```
public int sizeVertices() {
    return numVertices;
public int sizeEdges() {

    Equivale a contar todos los

    int numEdges=0;
                                        elementos true en la matriz.
    if (directed) {
        for (int i=0;i<numVertices;i++) {</pre>
             for (int j=0;j<numVertices;j++) {</pre>
                 if (matrix[i][j]!=false) numEdges++;
    } else {
        for (int i=0;i<numVertices;i++) {</pre>
             for (int j=i;j<numVertices;j++) {</pre>
                 if (matrix[i][j]!=false) numEdges++;
                             Si no es dirigido, como la matriz es
                             simétrica, sólo necesitaremos visitar una de
    return numEdges;
                             las dos partes divididas por la diagonal.
```

```
public int getOutDegree(int i) {
    if (!directed) {
        System.out.println("Graph non directed!!!");
        return 0;
    //checks if the vertex is right
    if (!checkVertex(i))
        throw new IllegalArgumentException("Nonexistent vertex " + i);
    int outdeg=0;
    for (int col=0;col<numVertices;col++) {</pre>
        if (matrix[i][col]!=false) outdeg++;
                                      Las filas representan los vértices de origen
    return outdeg;
                                      y las columnas los vértices destino
```

• Incrementamos 1 por cada columna cuyo índice tenga una arista con i, es decir, matrix[i,col].

```
public int getInDegree(int i) {
    if (!directed) {
        System.out.println("Graph non directed!!!");
        return 0;
    if (!checkVertex(i))
        throw new IllegalArgumentException("Nonexistent vertex " + i);
    int indeg=0;
    for (int row=0;row<numVertices;row++) {</pre>
        if (matrix[row][i]!=false) indeg++;
                                     Las filas representan los vértices de origen y
                                     las columnas los vértices destino
    return indeg;
```

• Incrementamos 1 por cada fila cuyo índice tenga una arista con i, es decir, matrix[row,i].

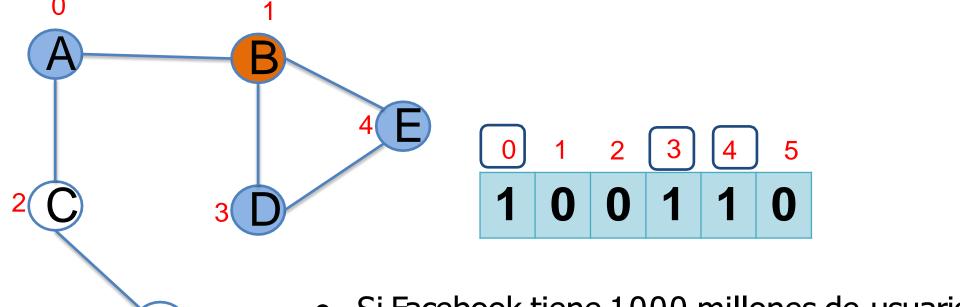
```
public int getDegree(int i) {
    if (!checkVertex(i))
        throw new IllegalArgumentException("Nonexistent vertex " + i);
    int degree=0;
    if (directed) degree=getInDegree(i)+getOutDegree(i);
    else {
        for (int row=0;row<numVertices;row++) {
            if (matrix[row][i]!=false) degree++;
        }
    }
}
return degree;
Siel grafo no es dirigido el grado será la suma del</pre>
```

Si el grafo no es dirigido, el grado será la suma del grado de entrada y el grado de salida.

En otro caso, bastará con que contemos las aristas de entrada en ese vértice. También se podría hacer contando las aristas de salida (pero nunca ambas).

```
//returns an array with the adjacent vertices for i
public int□ getAdjacents(int i) {
    if (!checkVertex(i))
            throw new IllegalArgumentException("Nonexistent vertex " + i);
    //obtains the number of adjacent vertices,
    //which will be the size of the array
    int numAdjacents=0;
    if (directed) numAdjacents=getOutDegree(i);
    else numAdjacents=getDegree(i);
    int[] adjacents=new int[numAdjacents];
    if (numAdjacents>0) {
        int j=0;
        //gets the edges (i,col) and saves col into adjacents
        for (int col=0; col<numVertices;col++) {</pre>
            if (matrix[i][col]!=null) {
                adjacents[j]=col;
                ]++;
    //return an array with the adjacent vertices of i
    return adjacents;
}
```

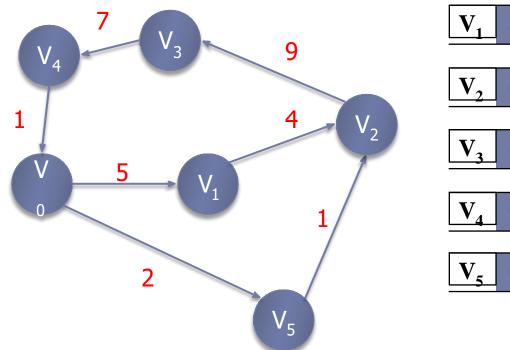
Matriz de Adyacencia - Conclusiones

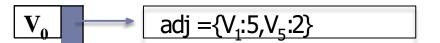


- Si Facebook tiene 1000 millones de usuarios (10⁹), las filas de la matriz de adyacencia son de dimensión 10⁹
- Si un usuario, B, tiene 1000 amigos, en su fila, habrá:
 - Número de 1s: 1000 = 1KB y
 - \circ Número de 0s: 10^9 1000 = 1GB

- La matriz de adyacencia consume memoria y la complejidad de las operaciones con la matriz es alta (por ejemplo, el método que muestra la matriz tiene complejidad cuadrática).
- Una lista de adyacencia sólo almacena la información para los aristas existentes, en lugar de almacenar todas las posibles combinaciones como ocurría en la matriz de adyacencias.
- Necesita menos espacio de memoria y su coste computacional es menor.

Lista adyacencias





$$V_1 = \{V_2:4\}$$

$$\mathbf{V}_{2} \longrightarrow \mathbf{adj} = \{\mathbf{V}_{3}:9\}$$

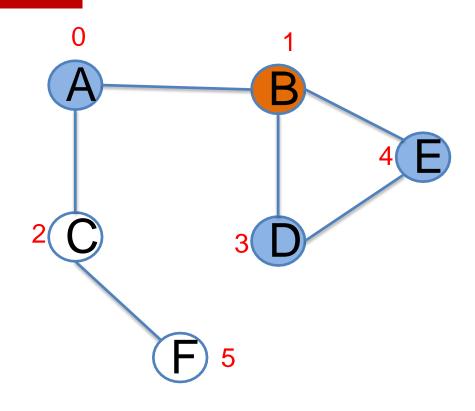
$$V_3 = \{V_4:7\}$$

$$\mathbf{V}_{4} \qquad \text{adj = } \{\mathbf{V}_{0}:\mathbf{1}\}$$

$$V_5 \qquad \text{adj =} \{V_2:1\}$$

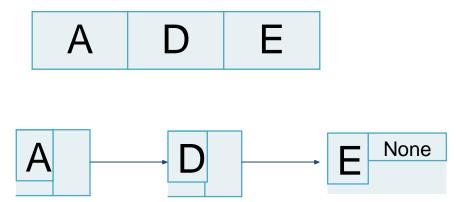
- Si tenemos un número fijo de vértices, el grafo se puede representar como un array de listas enlazadas.
- Cada posición del array representa un vértice y almacena la referencia a la lista de vértices adyacentes a dicho vértice (implementada como lista enlazada).
- ▶ Cada uno de los nodos almacenará la información sobre el vértice adyacente. Si el grafo es ponderado, también debería almacenar su peso.

Lista de Adyacencia

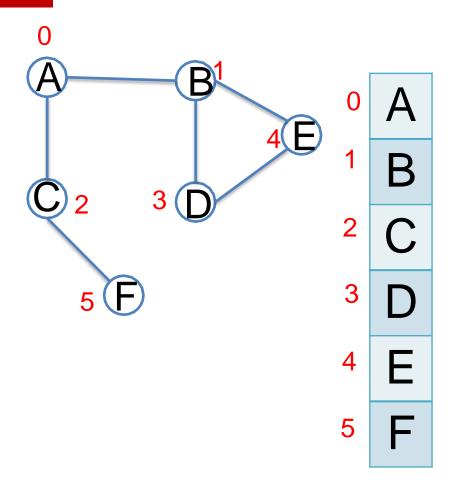


Para representar los <u>vecinos</u> <u>de B</u>, sería suficiente con almacenar sus índices en un lista de python o en una lista enlazada.

Vecinos de B:

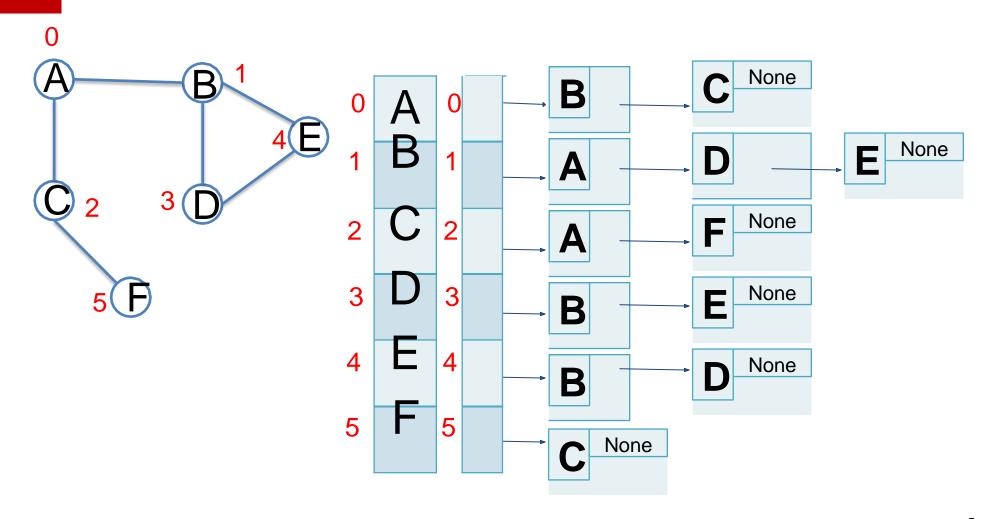


Lista de Adyacencia



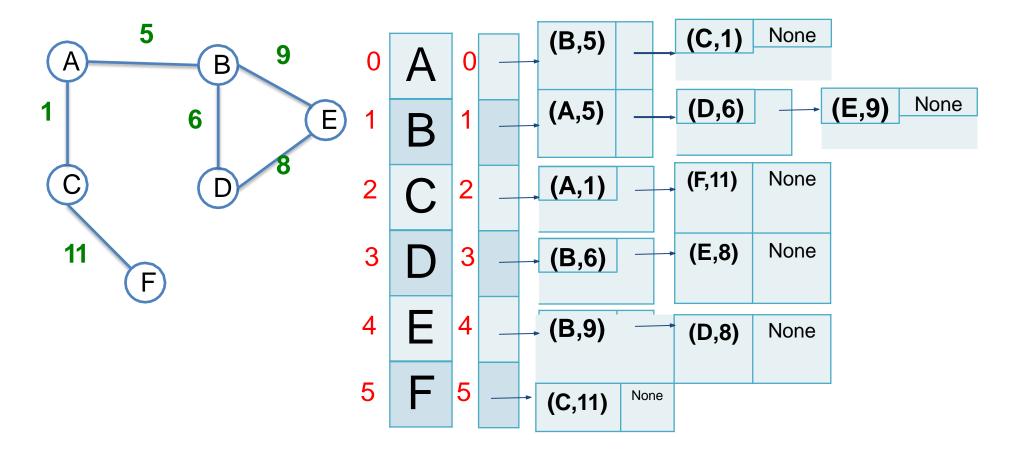
Seguiremos utilizando un array (lista de python) para almacenar los vértices

Lista de Adyacencia



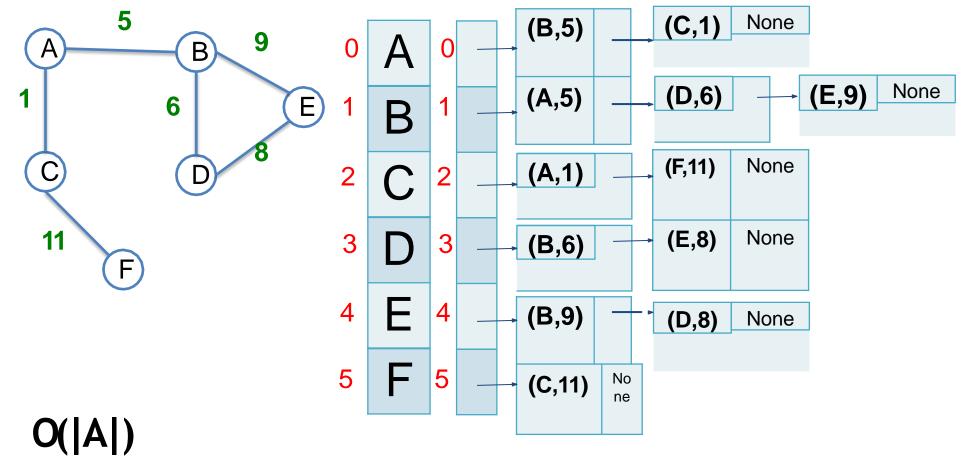
Usaremos un array de listas enlazadas para almacenar los vértices adyacentes a cada vértice.

Lista de Adyacencia (grafo ponderado)



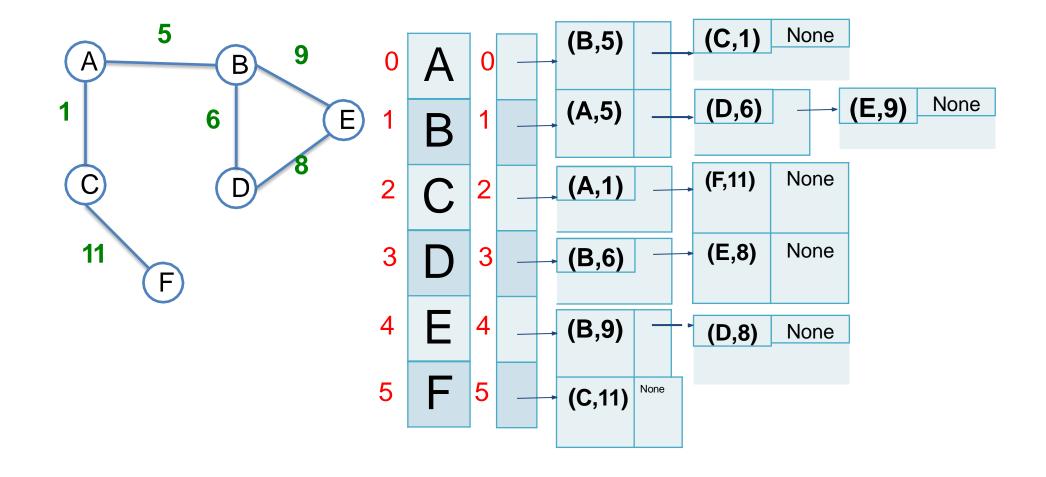
Cada vértice adyacente se representa como un par (v,w) donde v es el vértice adyacente, y el peso de la arista.

Lista de Adyacencia - Complejidad Espacial



Recuerda que el número máximo de aristas en un grafo simple es n(n-1)/2 (no dirigido) y n(n-1) (dirigido)

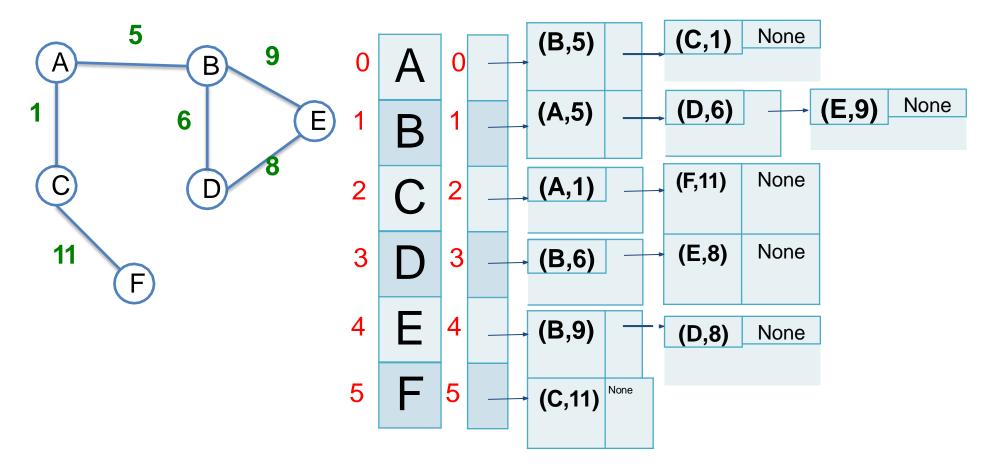
Lista de Adyacencia - Complejidad Espacial



Si el grafo es denso: $O(|A|) \rightarrow n^2$

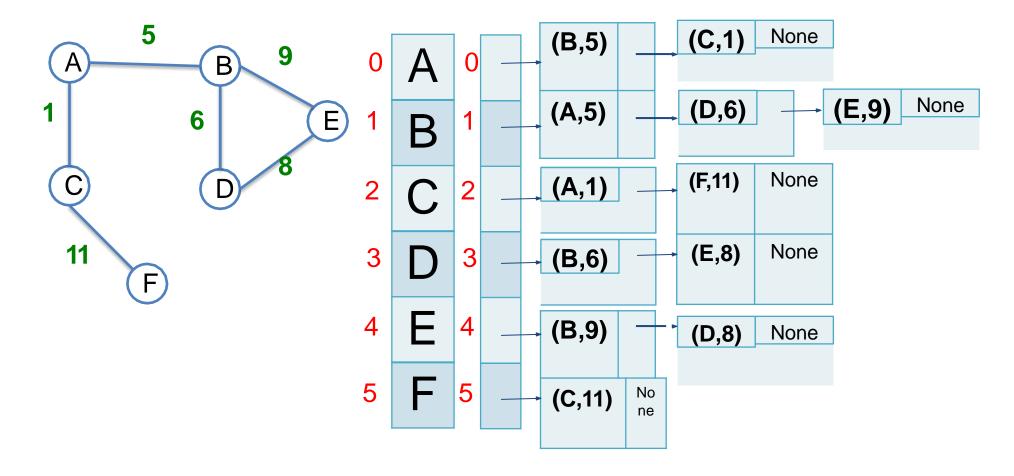
Si el grafo es disperso: O(|A|) -> n

Lista de Adyacencia - Complejidad Temporal



Obtener vértices adyacentes a E? O(n) (obtener su índice=4) y O(1) devolver la lista de adyacencia asociada al índice

Lista de Adyacencia - Complejidad Temporal



Comprobar si E y F son vecinos?
O(n) (obtener el índice de E) y O(n) comprobar si F está en la lista de adyacencia de E

Lista de Adyacencia - Conclusiones

- En términos de **complejidad temporal**, la lista de adyacencia y la matriz de adyacencia son similares. Para la mayorías de las operaciones, su complejidad es **O(n)** donde n=|**V**|.
- Sin embargo, en términos de complejidad espacial, la lista de adyacencia es una estructura más eficiente O(|A|).
- La mayoría de los grafos reales son escasos, y por tanto,
 |A|≈|V|=nPor tanto, la complejidad espacial será O(n).

```
import dlist.DListVertex;
public class GraphLAFull implements IGraph {
    int numVertices;
    int maxVertices;
    DListVertex[] vertices;
    boolean directed;
```

```
public GraphLAFull(int n, int max, boolean d) {
    if (max <= 0)
        throw new IllegalArgumentException("Negative maximum number of vertices!!!");
    if (n \le 0)
        throw new IllegalArgumentException("Negative number of vertices!!!.");
    if (n>max)
        throw new IllegalArgumentException("number of vertices can never "
                + "be greater than the maximum.");
    maxVertices=max;
    vertices=new DListVertex[maxVertices];
    numVertices=n;
    //creates each list
    for (int i=0; i<numVertices;i++) {</pre>
        vertices[i]=new DListVertex();
    directed=d;
```

Comprobamos que los índices son correctos.

```
public void addEdge(int i, int j, float w) {
    if (!checkVertex(i))
        throw new IllegalArgumentException("Nonexistent vertex " + i);
    if (!checkVertex(j))
        throw new IllegalArgumentException("Nonexistent vertex " + j);
        Tenemos que añadir el vértice j a la lista de
    vertices[i].addLast(j,w);
        vertices adyacentes del vértice i (que está
        almacenada en vertices[i]).
    if (!directed) vertices[j].addLast(i,w);
}

Si el grafo no es dirigido, deberemos también
    almacenar el vértice i como adyacente del vértice j
```

```
public void removeEdge(int i, int j) {
    if (!checkVertex(i))
        throw new IllegalArgumentException("Nonexistent vertex " + i);
    if (!checkVertex(j))
        throw new IllegalArgumentException("Nonexistent vertex " + j);
    int index=vertices[i].getIndexOf(j);
    vertices[i].removeAt(index);
    if (!directed) {
            index=vertices[j].getIndex0f(i);
            vertices[j].removeAt(index);
                           Si el grafo no es dirigido, deberemos también borrar
                           el vértice i como adyacente del vértice j
```

```
public boolean isEdge(int i, int j) {
    if (!checkVertex(i))
        throw new IllegalArgumentException("Nonexistent vertex " + i);
    if (!checkVertex(j))
        throw new IllegalArgumentException("Nonexistent vertex " + j);
    boolean result=vertices[i].contains(j);
    return result;
}
```

```
public int getOutDegree(int i) {
    if (!checkVertex(i))
        throw new IllegalArgumentException("Nonexistent vertex " + i);
    int outdegree=0;
    outdegree=vertices[i].getSize();
    return outdegree;
                                             Sólo para grafos
                                             dirigidos
public int getInDegree(int i) {
    if (!checkVertex(i))
        throw new IllegalArgumentException("Nonexistent vertex " + i);
    int indegree=0;
    for (int j=0; j<numVertices; j++) {</pre>
        if (vertices[j].contains(i)) indegree++;
    return indegree;
```

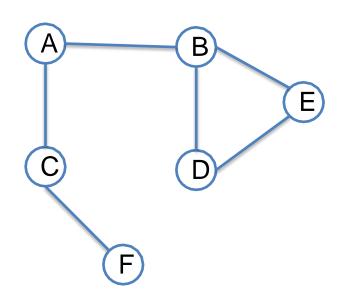
```
public int getDegree(int i) {
    int degree=0;
    if (directed) {
        degree=getOutDegree(i)+getInDegree(i);
    } else degree=vertices[i].getSize();
    return degree;
}
```

El grado de un vértice en un grafo dirigido es igual a la suma de su grado de entrada y de su grado de salida.

En un grafo no dirigido, es suficiente con obtener el número de vértices adyacentes a dicho vértice.

```
public int[] getAdjacents(int i) {
    if (!checkVertex(i))
        throw new IllegalArgumentException("Nonexistent vertex " + i);
    //gets the number of adjacent vertices
    int numAdj=vertices[i].getSize();
    //creates the array
    int[] adjVertices=new int[numAdj];
    //saves the adjacent vertices into the array
    for (int j=0; j<numAdj; j++) {</pre>
        adjVertices[j]=vertices[i].getVertexAt(j);
    //return the array with the adjacent vertices of i
    return adjVertices;
```

Diccionarios



Los vértices del grafo se van a representar como las claves (keys)del diccionario. En el diccionario, asociado a cada clave (vértice), se guarda la lista de sus vértices adyacentes graph ={ 'A':['B',C'],

'B':['A',D,E'],

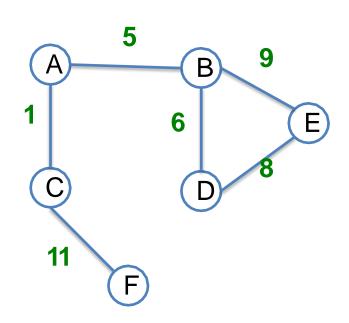
`C':['A','F'],

'D':['B",E'],

`E':['B','D'],

`F':['C']}

Diccionarios (grafos ponderados)



```
graph = {
'A':[('B',5),('C',1)],
'B':[('A',5),('D',6),('E',9)],
'C':[('A',1),('F',11)],
'D':[('B',6),('E',8)],
'E':[('B',9),('D',8)],
'F':[('C',11)] }
```

Implementar el TAD grafo