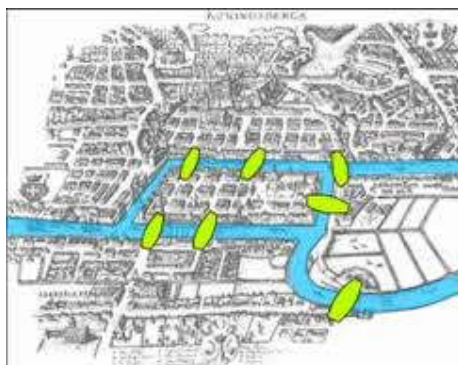
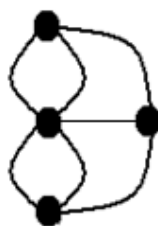


Grafos

El origen del concepto de grafo se remonta a la ciudad de Königsberg , en Prusia (actualmente, Kaliningrado, en Rusia). Por esa ciudad pasa el río Pregel y existen siete puentes, como muestra la figura. En épocas del matemático Euler, la gente del lugar se desafiaba a hacer un recorrido que permitiera atravesar cada puente una sola vez regresando al punto de partida.



Este problema, que tenía gran dificultad ya que nadie superaba el desafío, atrajo la atención de Euler, quien lo analizó empleando una técnica de graficado por la cual redujo a puntos la representación de las islas y las orillas, y a arcos los siete puentes. El grafico resulto del siguiente modo. A los puntos se los denomino ‘nodos’ o ‘vértices’. Los enlaces son ‘aristas’ o ‘arcos’.



El problema original se volvió entonces ahora es análogo al de intentar dibujar el grafo anterior partiendo de un vértice, sin levantar el lápiz, sin pasar dos veces por el mismo arco y terminando en el mismo vértice donde se había comenzado.

Para relater de forma somera la resolución de Euler, podemos decir que se puede considerar que en un grafico de este tipo hay nodos “intermedios” y de “inicio o finalización”. Los nodos intermedios tienen un número par de aristas incidentes en ellos (ya que “si se llega” a uno de ellos se debe “volver a salir”, puesto que por eso son intermedios). Se demuestra que, para que sea posible realizar el dibujo en las condiciones indicadas, si el inicio y la finalización se hace en nodos distintos, estos nodos, y solamente estos deben tener un número impar de aristas incidentes; el resto de los nodos debe tener “grado” par. Si lo que se quiere es comenzar y terminar en el mismo vértice, entonces todos los nodos deben tener un número par de aristas incidentes en ellos.

Analizando la situación de los puentes, se concluye que el problema de los puentes de Königsberg no tenía solución.

De este modo comenzó la teoría de Grafos.

Los grafos constituyen una muy útil herramienta matemática para modelizar situaciones referidas a cuestiones tan diferentes como mapas de interrelación de datos, carreteras, cañerías, circuitos eléctricos, diagrama de dependencia, etc.

Definición matemática de Grafo

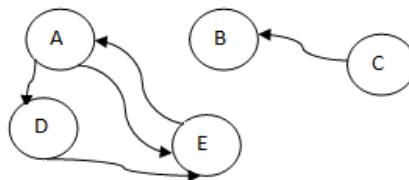
Los grafos son orientados (o dirigidos o digrafos) si las aristas (o arcos) que conectan sus vertices (también llamados nodos) están orientadas.

Los grafos son no orientados (o no dirigidos) si las aristas que conectan sus vertices no están orientadas.

Un grafo dirigido o digrafo o grafo orientado consiste en una dupla formada por un conjunto V de vértices o nodos del grafo, y un conjunto de pares ordenados A pertenecientes a $V \times V$.

En símbolos el grafo dirigido G es

$G = (V, A)$ donde A es un subconjunto de $V \times V$ Ejemplo:



$V = \{A, B, C, D, E\}$

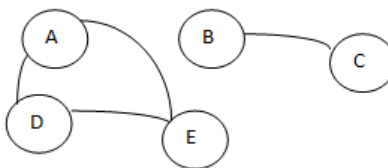
$A = \{ (A,D), (A,E), (E,A), (D,E), (C,B) \}$

Un grafo no dirigido (o no orientado) es una dupla formada por un conjunto V de vértices o nodos del grafo, y un conjunto de pares NO ordenados A (aristas no orientadas) pertenecientes a $V \times V$.

En símbolos el grafo no dirigido G es

$G = (V, A)$ donde A es un conjunto de pares no ordenados de $V \times V$

Ejemplo:



$V = \{A, B, C, D, E\}$

$A = \{ (A,D), (A,E), (D,E), (C,B) \}$ (pares no ordenados)

En ambos casos, si (v,w) pertenece a A se dice que w es adyacente a v .

Algunas definiciones más:

Camino: secuencia de vértices v_1, v_2, \dots, v_n tales que
 v_2 es adyacente a v_1
 v_3 es adyacente a v_2
...
 v_n es adyacente a v_{n-1}

Longitud de Camino: Cantidad de Aristas del camino.

Camino Abierto: Camino donde el vértice inicial y final difieren.

Camino Cerrado: Camino donde el vértice inicial y final coinciden.

Recorrido: Camino que no repite aristas.

Ciclo: camino que conteniendo vertices distintos, excepto el primero que coincide con el ultimo.

Subgrafo: dado el grafo $G = (V, A)$, un *subgrafo* de G es un grafo $G' = (V', A')$ que cumple V' es un subconjunto de V y A' es un subconjunto de A .

Grafo no dirigido conexo: un grafo no orientado es conexo sii para todo vértice del grafo hay un camino que lo conecte con otro vértice cualquiera del grafo

Grafo subyacente de un grafo dirigido: es el grafo no dirigido que se obtiene reemplazando cada arista (orientada) del mismo, por una arista no orientada.

Grafo dirigido fuertemente conexo: un grafo dirigido es fuertemente conexo sii entre cualquier par de vértices hay un camino que los une.

Grafo dirigido débilmente conexo: es aquel grafo dirigido que no es fuertemente conexo y cuyo grafo subyacente es conexo.

Árbol libre: es un grafo no dirigido conexo sin ciclos.

El TDA Grafo:

Es un TDA contenedor de un conjunto de datos llamados nodos y de un conjunto de aristas cada una de las cuales se determina mediante un par de nodos.

Crear grafo: esta primitiva genera un grafo vacío.

Precondición: -----

Poscondición: grafo generado vacío

Destruir grafo: esta primitiva destruye el grafo.

Precondición: que el grafo exista .

Poscondición: -----

Insertar nodo: esta primitiva inserta un nodo nuevo, recibido como argumento, en el grafo

Precondición: que el grafo exista y que el nodo no esté previamente

Poscondición: el grafo queda modificado por el agregado del nuevo nodo

Insertar arista: esta primitiva inserta una arista nueva, recibida como argumento, en el grafo

Precondición: que el grafo exista, que la arista no esté previamente y que existan en el grafo los nodos origen y destino de la arista.

Poscondición: el grafo queda modificado por el agregado de la nueva arista

Eliminar nodo: esta primitiva elimina un nodo, recibido como argumento, del grafo

Precondición: que el grafo exista y que el nodo a eliminar esté en él y no tenga aristas incidentes en él.

Poscondición: el grafo queda modificado por la eliminación del nodo

Eliminar arista: esta primitiva elimina una arista, recibida como argumento, del grafo

Precondición: que el grafo exista y la arista estén en él.

Poscondición: el grafo queda modificado por la eliminación de la arista

Existe arista: esta primitiva recibe una arista y retorna un valor lógico indicando si la arista existe en el grafo

Precondición: que el grafo exista

Poscondición: -----

Existe nodo: esta primitiva recibe una arista y retorna un valor lógico indicando si el nodo existe en el grafo.

Precondición: que el grafo exista

Poscondición: -----

Pueden considerarse también las operaciones correspondientes a los recorridos como básicas en el TDA (o bien plantear las utilizando iteradores para navegar dentro del contenedor).

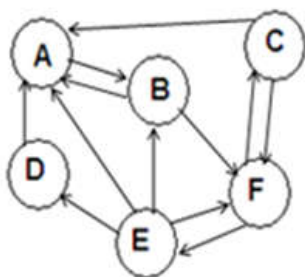
Estructuras de datos utilizadas para implementar grafos

Como sucede en la mayoría de los TDA Contenedores, las propuestas básicas de implementaciónn almacenan los datos de forma contigua (por ejemplo, usando arrays, matrices o tablas) o de forma dispersa (utilizando listas con punteros).

Entonces, la primera alternativa es utilizar una Matriz de Adyacencia, la cual, para N nodos es de $N \times N$.

Si M es la matriz de adyacencia del grafo G entonces verifica que $M[i][j]$ es true (o 1) si la arista (i,j) pertenece al grafo, o false (o bien 0) si no pertenece. Adicionalmente, en alguna estructura se puede almacenar, si es necesario, información adicional sobre los vertices (identificadores o etiquetas, y quizás valores asociados)

Ejemplo:

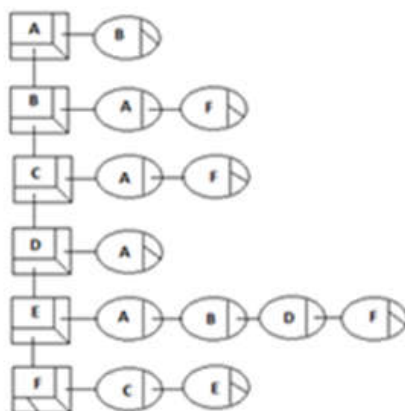
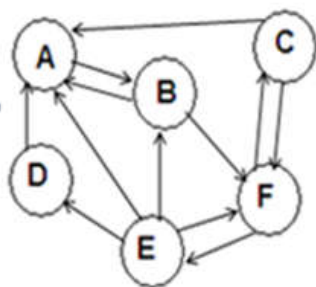


	A	B	C	D	E	F
A	0	1	0	0	0	0
B	1	0	0	0	0	1
C	1	0	0	0	0	1
D	1	0	0	0	0	0
E	1	1	0	1	0	1
F	0	0	1	0	1	0

Otra posibilidad es usar listas de adyacencia. En este caso, la representación es mediante una lista de listas. De este modo, se tiene una lista con nodos; cada uno de ellos conteniendo la información sobre un vértice y un puntero a una lista ligada a los vértices adyacentes al vértice indicado.

Ejemplo:

1.



Recorridos de un grafo dirigido:

Los dos recorridos básicos en los grafos dirigidos son en profundidad y en anchura. Cada uno de ellos procesa o visita cada vértice del grafo, visitando una y solo una vez a cada uno.

Recorrido en profundidad (*depth first search* o ‘busqueda en profundidad’):

Consiste en hacer lo siguiente:

Inicialmente se marcan todos los nodos como “no visitados”.

Luego, mientras haya vértices no visitados, se toma el primero de ellos, y se lo visita y marca (como visitado) para luego pasar al primer adyacente de este vértice que se haya no visitado y proceder del mismo modo, pasando luego al primer adyacente del adyacente que no haya sido visitado, y así hasta llegar a un nodo que no tenga adyacentes no visitados.

El recorrido concluye con la visita de todos los nodos, no pasando por ninguno de ellos más de una vez.

Para implementar el algoritmo suele usarse una pila, o bien un algoritmo recursivo.

Puede, además, numerarse los vértices a medida que se visitan, tal como se lleva a cabo en los algoritmos que se describen a continuación.

Recorrido en Profundidad recursivo con numeración de vértices

//Se considera que cada vértice tiene un atributo número que almacena el orden en que se lo visitó.

//Inicialmente todos los atributos numero valen 0. Eso significa que no fueron visitados.

//Índice es una variable que permite realizar el conteo de los nodos visitados para asignar el valor a

// cada atributo número de cada vértice

DFSRecursivoCon Numeracion

```
{
Para cada vértice v del grafo G
    {
        Asignar 0 a número de v;
        indice= 0;
    }

Para cada vértice v
    {
        Si (numero de v es 0) entonces
            {Numerar (v, indice);} //indice pasa por referencia; su valor se actualiza entre llamadas
        }
    }
}
```

Numerar (u:vértice; var nd: entero)

```
{
    nd = nd+1;
    Asignar nd al atributo número del vértice u;
    Para cada vértice w adyacente a u
        {
            Si (numero de w es 0 ) entonces
                {Numerar (w, nd);}
            }
        }
}
```

Análisis del coste temporal:

El coste depende de las estructuras que se usen para almacenar el grafo.

Si tenemos un grafo G con una cantidad v de vértices y a aristas y tenemos una implementación de matriz de adyacencia, observamos que:

El ciclo que inicializa numero de cada vértice marcándolo como no visitado es $O(v)$.

El ciclo que revisa si cada vértice ha sido o no visitado para invocar a Num tiene $O(v)$ iteraciones.

El coste del bloque depende del coste de la function Numerar.

Numerar comienza con una secuencia de sentencias de coste $O(1)$. Luego hay un ciclo que analiza cada adyacente a u para invocar a Numerar si ese adyacente no ha sido visitado.

Para determinar los adyacentes a u si tenemos implementación en Matriz de Adyacencia hay que recorrer la fila correspondiente al vértice u en la matriz. Esto involucra $O(v)$ iteraciones, siendo v el número total de veces que se ejecuta Numerar. Entonces, el coste de la búsqueda en profundidad para el caso de una implementación con matriz de adyacencia pertenece a $O(v^2)$.

Si se ha implementado el grafo con listas de adyacencia,

Inicializar numero de cada vértice en 0 es $O(v)$.

Numerar se invoca para cada vértice, tantas veces como adyacentes tenga el vértice (pero hay que considerar que los vertices procesados por ser adyacentes a un vértice no son después procesados de nuevo). Los adyacentes a un vértice se determinan recorriendo la lista de vertices adyacentes. El proceso Numerar realizará el recorrido de todas las listas de nodos ayacentes. El costo se expresa por tanto como $O(a)$.

Entonces tenemos que el coste se puede expresar como $O(v) + O(a)$, lo que generalmente se indica como $O(v+a)$.

Recorrido en Profundidad iterativo con numeración de vértices:

//Como en el algoritmo anterior, se considera que cada vértice tiene un atributo número

//que almacena el orden en que se lo visitó. Inicialmente todos los atributos numero valen 0,

//lo cual indica que no fueron visitados

//Índice es una variable que permite realizar el conteo de los nodos visitados para asignar el valor a

// cada atributo número de cada vértice

// Este algoritmo utiliza una pila inicialmente vacía

DFSIterativoConNumeracion

```
{
    Inicializar pila p a vacío;
    indice = 0;
    Para cada vértice v
        {
            Asignar 0 a numero de v;
        }
    Apilar v en pila p;
    Mientras pila p no vacía
    {
```

```

Desapilar p en u;    //u es una variable tipo vértice
Si (el numero de u es 0)
{
    índice= índice +1;
    asignar índice a numero de u;
    Para cada w adyacente a u
        Si (numero de w es 0)
            {
                Apilar w en pila p;
            }
}
}

```

Análisis del coste temporal:

Si se trabaja con un grafo G con una cantidad v de vértices y a aristas y con una implementación de matriz de adyacencia, sucede lo siguiente:

Considerando $O(1)$ el coste de apilar y desapilar, el primer ciclo es $O(v)$.

Luego se tiene un ciclo que se ejecuta mientras no se haya vaciado la pila. En ese ciclo, si el vértice desapilado no se ha visitado (lo cual sucederá v veces en total), se lo numera y se analizan sus adyacentes (a lo sumo $v-1$) para apilarlos si no han sido visitados aún.

De lo cual se deduce un coste temporal $O(v^2)$ para el recorrido con la implementación indicada.

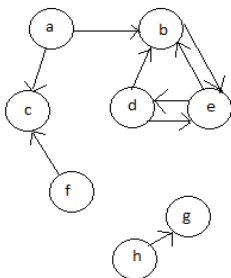
En cambio, con una implementación con listas de adyacencia,

El primer ciclo tiene coste $O(v)$.

Luego se tiene, para cada vértice desapilado y con numeración 0 (es decir, no visitado), el apilamiento de sus adyacentes. En total esto se realiza $O(a)$ de veces.

Entonces, el coste temporal total del recorrido en profundidad iterativo con implementación en listas de adyacencia es $O(v + a)$.

Ejemplo: para el siguiente grafo se muestra el recorrido en profundidad:



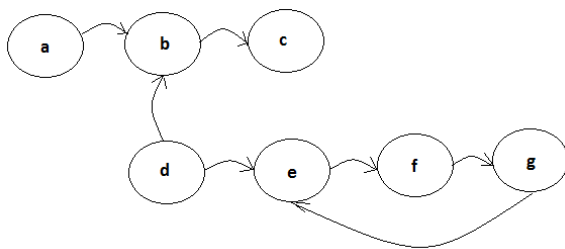
a	b	e	d	c	f	g	h
---	---	---	---	---	---	---	---

1	2	3	4	5	6	7	8
---	---	---	---	---	---	---	---

Test de aciclidad: (aplicación del recorrido en profundidad).

A menudo es necesario determinar si un grafo tiene o no ciclos (de no tenerlos se lo denomina acíclico). Para averiguarlo es posible usar el recorrido en profundidad, y detectar, si hay “aristas de retroceso”, que conecten un vértice con otro que aparece como adyacente, pero que ya ha sido visitado antes en el recorrido parcial que se está realizando (es decir, la cuestión no es si se nos encontramos con un vértice que ya ha sido visitado, sino si ese vértice ya visitado lo es en el recorrido parcial que se está realizando).

Ejemplo:



En la figura se observa que, partiendo del vértice a se visita a-b-c. Ese es el primer “recorrido parcial”. Luego se comienza con d un nuevo recorrido parcial, porque es el siguiente vértice no visitado. Se visita d. Los adyacentes a d son b y e. El vértice b ya ha sido visitado, pero esto no significa que exista un ciclo. De d se pasa y visita e, luego se visita f y luego g. El vértice e es adyacente a g y ya ha sido visitado, pero lo ha sido en este recorrido parcial. Eso es lo que indica que hay un ciclo. Por lo cual se deberá llevar un “doble registro” de vértices visitados: para todo el grafo y para cada recorrido parcial. La información de lo que sucede en cada recorrido parcial es la que permite detectar ciclos.

El algoritmo puede plantearse de la siguiente forma:

```

//La variable lógica b almacena el valor a retornar
//Este algoritmo utiliza para cada vértice un atributo que indica si fue visitado alguna vez (en
//cualquiera de los recorridos realizados, es una “marca general”)
//Y usa otro atributo llamado en_recorrido_actual para indicar si fue visitado en el recorridoparcial
que se esté realizando
//Luego de terminado cada recorrido parcial, el atributo en_recorrido_actual se reinicializa
  
```

Test_Aciclidad

```

{
  Inicializar b en falso; //valor lógico a retornar
  Para cada vértice v
  {
    Marcar v como no visitado;
    Asignar falso a atributo en_recorrido_actual de v;
  }
}
  
```

```
Para cada vértice v
{
    Si (b== falso) y (v no visitado)
    {
        Analizar_Recorrido_Parcial(v, b); // b pasa por referencia
    }
    Retornar b;
}
```

```
Analizar_Recorrido_Parcial ( u :vértice; var b: bool )
{
    Marcar u como visitado;
    Asignar verdadero a atributo en_recorrido_actual de u;
    Para cada w adyacente a u
    {
        Si (b==falso)
        {
            Si (w visitado) y (en_recorrido_actual de w == verdadero)
            {asignar verdadero a b;}
            Si (w no visitado )
            { Analizar_Recorrido_Parcial ( w,b ); }
        }
    }
    Asignar falso a atributo en_recorrido_actual de u;
}
```

Análisis del coste temporal:

Los costes son los mismos que se han discutido para el recorrido en profundidad, según la implementación que se haya hecho para el grafo.

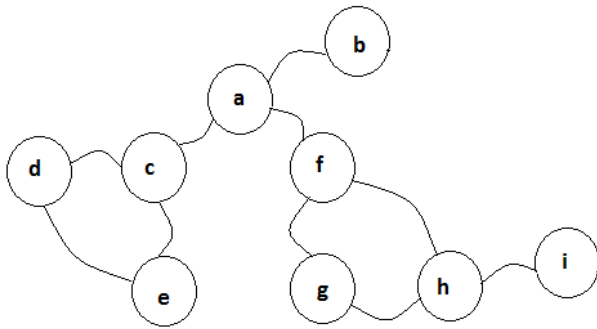
Puntos de articulación de un grafo: (aplicación del recorrido en profundidad).

Otra información que puede obtenerse de un grafo a través del recorrido en profundidad, es el conjunto de puntos de articulación.

Un punto de articulación (o vértice de corte) de un grafo no dirigido y conexo es un vértice que verifica que al ser eliminado del grafo éste deja de ser conexo.

Ejemplo:

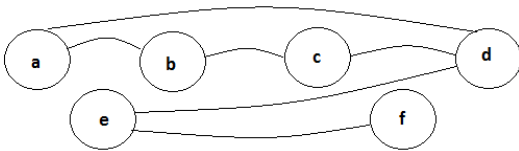
En la siguiente figura se puede observar que si se elimina alguno de estos vértices: a, c, f, h, el grafo deja de ser conexo. Esos vértices son puntos de articulación o vértices de corte.



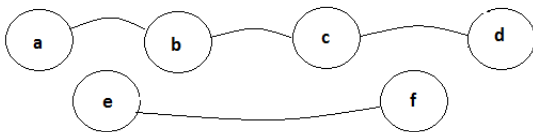
Si un grafo no tiene puntos de articulación significa que para todo par de vértices existe más de un camino que los enlaza. Entonces, si un vértice del grafo “se cayera” (es decir, si hubiera que eliminarlo del grafo) el grafo permanecería aún conexo. A los grafos que no tienen puntos de articulación se los llama biconexos.

Vamos a obtener los puntos de articulación de un grafo no dirigido aplicando un recorrido en profundidad. Para esto consideremos que cuando realizamos un recorrido de ese tipo, cada recorrido parcial (es decir el que comenzamos con un vértice en particular) es un “árbol libre” tal como lo hemos definido al principio. El conjunto de estos árboles obtenidos cada uno por un recorrido parcial se denomina “bosque”.

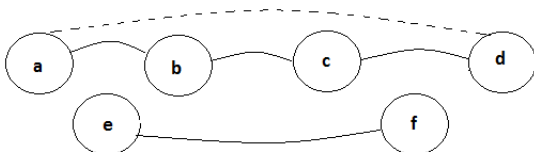
En el transcurso del recorrido, algunas aristas del grafo no aparecerán en el árbol de un recorrido parcial por ser “aristas de retroceso”. Ejemplo: en el grafo de la siguiente figura,



Este es el bosque del recorrido en profundidad. Tiene dos árboles.



Se ha marcado en línea punteada la arista de retroceso.



Ahora bien, si consideramos un grafo no dirigido y conexo, el bosque generado en un DFS, y determinadas las aristas de retroceso,

- Si r es raíz es el árbol DFS y tiene más de un hijo en el árbol, entonces r es un punto de articulación.
- Para todo vértice u que no sea raíz en árbol DFS, u es punto de articulación si al eliminarlo del árbol resulta imposible “volver” desde alguno de sus descendientes (en el árbol) hasta alguno de los antecesores de u .

Para analizar el caso de los vértices que no son raíz, queremos saber para cada uno cuan atrás podemos llegar desde él. Consideraremos la función bajo, que tiene el menor número entre: el número que le asignamos al realizar el recorrido en profundidad, el número de algún vértice alcanzado desde él mediante una arista de retroceso, y el valor bajo alcanzado por alguno de sus hijos.

$\text{bajo}(u)$ = mínimo número asignado en el recorrido en profundidad considerando estos valores:

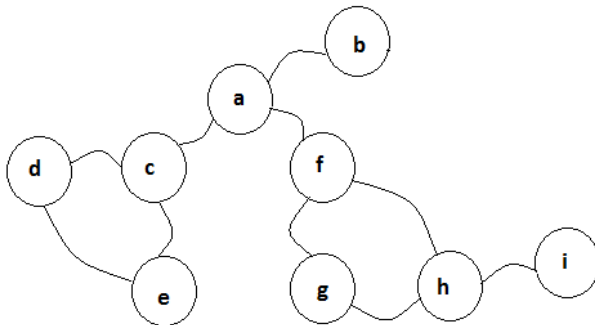
- Número asignado al vértice u en el recorrido en profundidad
- Número asignado en el recorrido en profundidad a cada nodo w al que se accede desde u por una arista de retroceso
- $\text{bajo}(x)$ para todo x hijo de u en el árbol que se genera en el recorrido.

Entonces el vértice u , si no es raíz, es punto de articulación si y sólo si tiene al menos un hijo x tal que $\text{bajo}(x) \geq \text{número asignado al vértice } u \text{ en el recorrido en profundidad}$.

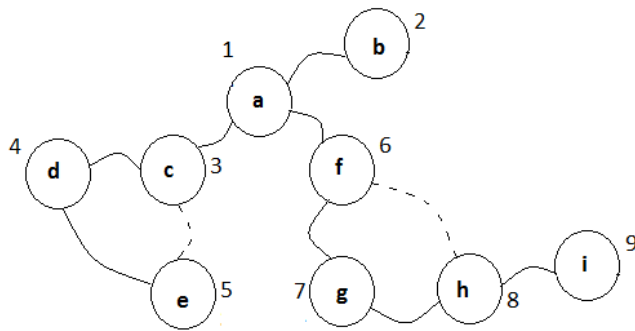
Esto es así porque significa que ese vértice x “necesita” de u para acceder a otros de numeración más baja. Si u se eliminara del grafo, dejaría de ser conexo.

A continuación se muestra el proceso de obtención de los puntos de articulación:

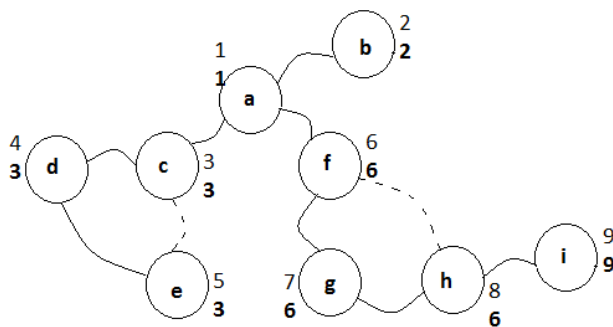
Partimos de



Numeramos cada vértice realizando el recorrido en profundidad y graficamos el árbol obtenido con el recorrido. En líneas punteadas las aristas de retroceso.



En la siguiente figura se ha colocado el valor de bajo(v) junto a cada vértice, debajo del número asignado en el recorrido.



Analizamos lo que sucede con cada vértice:

El vértice a es punto de articulación por ser raíz del árbol de recorrido.

Para los vértices no raíz consideremos lo que sucede con la siguiente tabla:

Vértice	Hijos
b (numeración: 2)	----
c (numeración: 3)	d (bajo: 3)
d (numeración: 4)	e (bajo: 3)
e (numeración: 5)	---
f (numeración: 6)	g (bajo: 6)
g (numeración: 7)	h (bajo: 6)
h (numeración: 8)	i (bajo: 9)
i (numeración: 9)	---

Observamos que los vértices **c**, **f** y **h** tienen hijos cuyo valor de bajo es mayor o igual que el número del vértice padre. Son puntos de articulación.

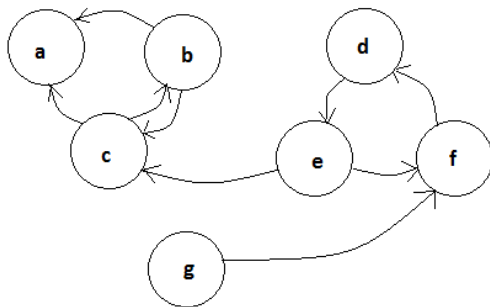
Entonces, tal como habíamos observado en el gráfico al principio, los puntos de articulación o vértices de corte de este grafo son: a, c, f, h.

Obtención de las componentes fuertemente conexas en un grafo dirigido: (aplicación del recorrido en profundidad).

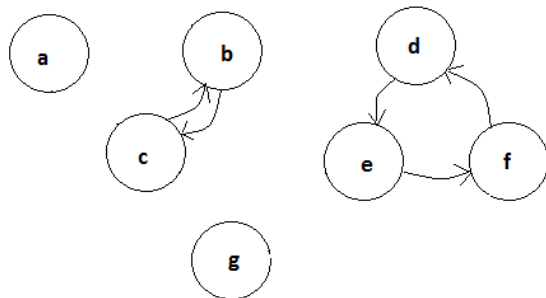
La búsqueda en profundidad también puede usarse para determinar con eficiencia las componentes fuertemente conexas de un conjunto.

Una componente fuertemente conexas de un grafo dirigido es un conjunto maximal de vértices en el que existe camino entre cualquier par de vértices.

Ejemplo: en la siguiente figura



Se muestran las componentes conexas; se han mantenido los enlaces entre vértices de cada componente:



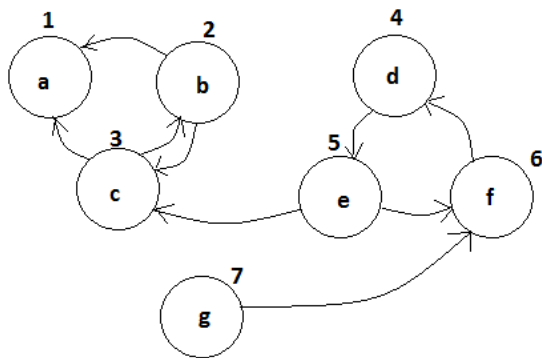
El algoritmo para hallar las componentes conexas de un grafo dirigido G es el siguiente:

- Realizar un recorrido en profundidad con numeración de vértices en G
- Construir un grafo dirigido nuevo G' , invirtiendo las direcciones de todas las aristas de G .
- Realizar un recorrido en profundidad en G' , partiendo del vértice con numeración más alta, de acuerdo con la numeración asignada en primer paso 1. Cuando termina un recorrido parcial tomando los vértices de modo decreciente, se continua con el siguiente vértice con numeración más alta.

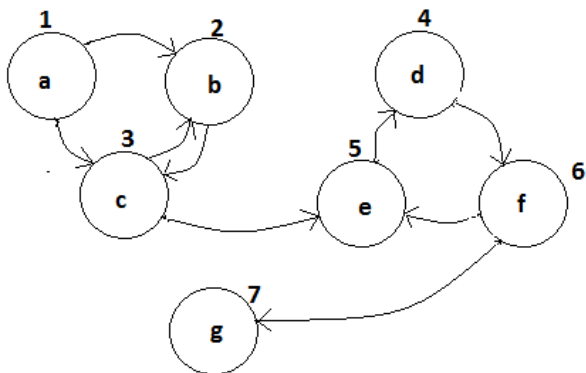
Al finalizar el recorrido, cada árbol correspondiente a un recorrido parcial es una componente fuertemente conexas del grafo.

Ejemplo:

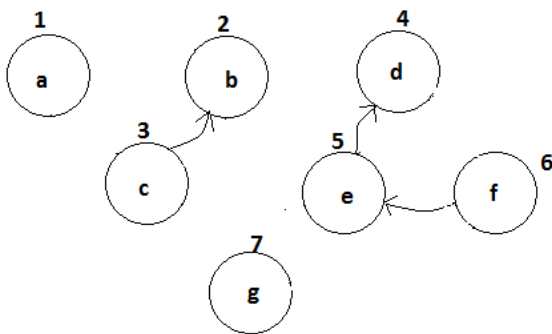
Primero, asignamos la numeración con el recorrido en profundidad en G.



Luego invertimos el sentido de las aristas y obtenemos G'.



Recorremos G' en profundidad, comenzando por el vértice 7 y terminando en el 1. Obtenemos:



Recorrido en anchura (breadth first search, o 'búsqueda primero en anchura'):

Este recorrido marca inicialmente todos los nodos como no visitados, y luego, para cada nodo no visitado, lo visita y marca, procediendo luego a visitar cada uno de sus adyacentes no visitados antes, luego de lo cual se continúa con los adyacentes del primer adyacente, los del segundo adyacente, etc., hasta agotar el conjunto de vértices del conjunto.

Se presenta a continuación un algoritmo que recorre en anchura numerando los vértices en el orden en que se visitan.

//Se considera que cada vértice tiene un atributo número //que almacena el orden en que se lo visitó.

//Inicialmente todos los atributos numero valen 0, lo cual indica que no fueron visitados

//Índice es una variable que permite realizar el conteo de los nodos visitados para asignar el valor a

// cada atributo número de cada vértice

// Este algoritmo utiliza una cola inicialmente vacía

BFSIterativoConNumeracion

```
{
Incializar cola c a vacío;
índice=0;
Para cada vértice v
    {
        Asignar 0 a numero de v
    }

Mientras exista un vértice v tal que numero de v sea 0
    {
        índice=índice+1;
        asignar índice a numero de v;
        Acolar v en cola c;
        Mientras cola no vacía
            {
                Desacolar en v;
                Para cada vértice u adyacente a v
                    {
                        Si (numero de u es 0)
                            {
                                índice=índice+1;
                                asignar índice a número de u;
                                Acolar u en cola c;
                            }
                    }
            }
    }
}
```

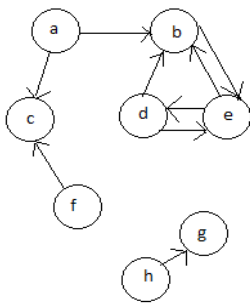
Análisis del coste temporal:

El coste depende de cómo se haya implementado el grafo.

Si se lo implementó con una matriz de adyacencia, tenemos lo siguiente: el coste de marcar cada nodo como no visitado es $O(v)$. Luego hay un ciclo de $O(v)$ iteraciones, en cada una de las cuales se analizan los adyacentes del vértice que se desacola (como la implementación es con matriz de adyacencia, esto tiene un coste $O(v)$). Por lo cual el coste del recorrido en anchura con implementación de matriz de adyacencia es $O(v^2)$.

Si la implementación es con listas de adyacencia, hay que considerar el coste de marcar cada nodo como no visitado, que es $O(v)$, y luego observar que en total tantas veces como aristas haya se va a analizar si el vértice adyacente al que se ha desacolado se ha visitado, para acolarlo si aún no se lo visitó. Entonces, el coste es $O(v) + O(a)$, indicado como $O(v+a)$.

Ejemplo de recorrido en anchura: se muestra el recorrido en anchura del siguiente grafo.



a	b	c	e	d	f	g	h
1	2	3	4	5	6	7	8

Recorridos topológicos:

Los recorridos topológicos son proceso de asignación de un orden lineal a los vértices de un grafo dirigido acíclico, de modo que se respeten las precedencias indicadas por las relaciones de adyacencia.

Estos recorridos solo se aplican a grafos dirigidos acíclicos, y permiten linealizar un grafo, es decir, recorrer los vértices del mismo respetando las precedencias.

Pueden plantearse recorridos topológicos como variantes del recorrido en profundidad y del recorrido en anchura.

Recorrido topológico en profundidad

//devuelve una lista lis en la cual quedan insertados los nodos según fueron visitados en el recorrido)
TopologicoDFS

```

{
  inicializar lista de salida lis a vacío; //lista vacía
  para cada vértice v
  {
    marcarlo como no visitado
  }
}
    
```

```

    }
    para cada vértice v
    {
        si (v no visitado) entonces
        {
            Rec (v, lis) //lis pasa por referencia
        }
        retornar lis;
    }
}

```

```

Rec ( v: vertice , var lis: lista)
{
    marcar v como visitado
    para cada vértice w adyacente a v
    {
        si (w no visitado) entonces
        {
            Rec(w, lis)
        }
        insertar v al frente en lista lis
    }
}

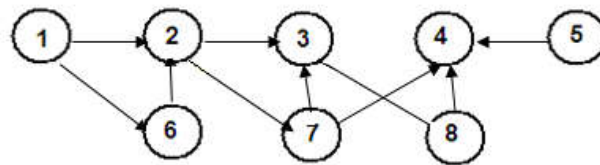
```

Análisis de coste temporal:

Análogo al del recorrido en profundidad habitual, según la implementación usada.

Ejemplo:

Consideremos el siguiente grafo dirigido y acíclico:



Según el algoritmo anterior,

```

{
Rec(1)
Adyacentes(1):2, 6
{
    Rec(2)

```

```

    Adjacentes(2):3, 7
    {
        Rec(3)
        Adjacentes(3):8
        {
            Rec(8)
            Adjacentes(8):4
            {
                Rec(4)
                Adjacentes(4):-
                Insertar 4 al frente en la lista de salida
            }
            Insertar 8 al frente en la lista de salida
        }
        Insertar 3 al frente en la lista de salida
    }
    {
        Rec(7)
        Adjacentes(7): 3, 4 (ambos ya visitados)
        Insertar 7 al frente en la lista de salida
    }
    Insertar 2 al frente en la lista de salida
}
{
    Rec(6)
    Adjacentes(6): 2 (ya visitado)
    Insertar 6 al frente en la lista de salida
}
Insertar 1 al frente en la lista de salida
}
{
    Rec(5)
    Adjacentes(5): 4 (ya visitado)
    Insertar 5 al frente en la lista de salida
}

```

De este modo, la lista generada que respeta las precedencias entre los nodos con un recorrido en profundidad es:

5, 1, 6, 2, 7, 3, 8, 4

Recorrido topológico en anchura

Se comienza calculando, para cada vértice, el grado de entrada del mismo (número de aristas que finalizan en el vértice).

Los vértices de grado de entrada 0 se acolan (tiene que haber vértices con grado de entrada 0; de no ser así, el grafo tendría ciclos).

Luego se procesa la cola del siguiente modo: mientras no esté vacía se desacola y visita un vértice (y se dirigirá a la salida, si queremos el listado respetando precedencias) y para cada vértice adyacente al mismo se decrementa en 1 el grado de entrada del mismo (como si se retirara el vértice desacolado del grafo). Cada vértice que llegue a 0 por este decremento, se acola.

// Este algoritmo genera una lista de salida lis, con todos los vértices del grafo

// La tabla T almacena para cada vértice su grado de entrada. Se va actualizando durante el proceso.

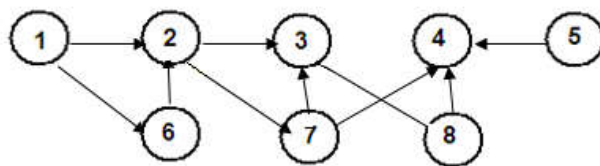
TopologicoBFS

```
{
  inicializar table T;
  Inicializar cola c a vacío; // cola para vértices del grafo; se inicializa vacía.
  Inicializar lista lis a vacío, lista de salida;
  Para cada vértice v
  {
    Almacenar el grado de entrada de v en la table T;
    Si (grado de entrada de v es 0)
    {
      Acolar v en cola c;
    }
  }
  Mientras (cola c no vacía)
  {
    Desacolar en x;
    Insertar x al final en lista;
    Para cada vértice w adyacente a v
    {
      Decrementar en 1 el grado de w registrado en la tabla T;
      si (grado de w registrado en tabla T es 0)
      {
        Acolar w en cola c;
      }
    }
  }
  Retornar lista lis;
}
```

Análisis de coste temporal: es análogo al de los algoritmos que aplican recorridos y que hemos considerado antes.

Ejemplo de recorrido topológico en anchura:

Para el grafo anterior,



La tabla de grados de entrada es:

vértice	1	2	3	4	5	6	7	8
Grado entrada	0	2	2	3	0	1	1	1

Se acolan el 1 y el 5

Cola : 1 – 5

Se desacola y procesa el 1

Se decrementan los adyacentes a 1, que son 2 y 6.

vértice	1	2	3	4	5	6	7	8
Grado entrada	0	1	2	3	0	0	1	1

6 ahora vale 0 y se acola.

Cola: 5 – 6

Se desacola y procesa 5.

Se decrementan los adyacentes a 5, que es 4.

vértice	1	2	3	4	5	6	7	8
Grado entrada	0	1	2	2	0	0	1	1

Cola: 6

Se desacola y procesa 6.

Se decrementa el adyacente a 6, que es 2.

vértice	1	2	3	4	5	6	7	8
Grado entrada	0	0	2	2	0	0	1	1

2 vale ahora 0 y se acola.

Cola: 2

Se desacola y procesa 2.

Se decrementan los adyacentes a 2, que son 3 y 7.

vértice	1	2	3	4	5	6	7	8
Grado entrada	0	0	1	2	0	0	0	1

7 llega a 0, se acola.

Cola: 7

Se desacola y procesa 7

Se decrementan los adyacentes a 7, que son 3 y 4

vértice	1	2	3	4	5	6	7	8
Grado entrada	0	0	0	1	0	0	0	1

3 llega a 0, se acola

Cola: 3

Desacolar 3

Decrementar el adyacente a 3, que es 8

vértice	1	2	3	4	5	6	7	8
Grado entrada	0	0	0	1	0	0	0	0

8 vale 0, se acola

Cola: 8

Desacolar 8

Decrementar el adyacente a 8, que es 4

vértice	1	2	3	4	5	6	7	8
Grado entrada	0	0	0	0	0	0	0	0

4 se acola porque llegó a 0.

Cola: 4

Desacolar y procesar 4, que no tiene adyacentes.

La linealización del grafo da por resultado esta lista (se han indicado las precedencias con arcos)

1-5- 6- 2- 7- 3- 8- 4

Problemas sobre caminos en un grafo:

Hay un gran número de problemas sobre caminos en grafos dirigidos y en grafos no dirigidos. Algunos plantean determinar si dado un par de vértices, o todos los pares de vértices, hay o no camino entre ellos.

Otros trabajan sobre grafos con aristas o con vértices ponderados. Una ponderación es un valor asociado a una arista o a un vértice, o a ambos.

En esta sección trataremos algunos de los problemas más comunes sobre caminos en grafos con aristas ponderadas.

Problema de los caminos más cortos con un solo origen

Dado un grafo dirigido $G = (V, A)$, en el cual cada arco tiene asociado un costo no negativo, y donde un vértice se considera como origen, el problema de los "caminos más cortos con un solo

origen" consiste en determinar el costo del camino más corto desde el vértice considerado origen a todos los otros vértices de V .

Este es uno de los problemas más comunes que se plantean para los grafos dirigidos con aristas ponderadas (es decir con peso en las aristas). La 'longitud' o 'costo' de un camino es la sumatoria de los pesos de las aristas que lo conforman.

El modelo de grafo dirigido con aristas ponderadas no negativas puede, por ejemplo, corresponder a un mapa de vuelos en el cual cada vértice represente una ciudad, y cada arista (v,w) una ruta aérea de la ciudad v a la ciudad w .

Algoritmo de Dijkstra para el cálculo de los caminos mínimos:

Desarrollado por Dijkstra en 1959 (acá la primera publicación sobre el tema <http://wwwm3.ma.tum.de/foswiki/pub/MN0506/WebHome/dijkstra.pdf>)

Este algoritmo se caracteriza por usar una estrategia llamada *greedy*, voraz, o ávida. Se exige, para aplicar esta estrategia en este problema que el peso de las aristas sea no negativo.

La estrategia greedy trata de optimizar (alcanzar el máximo o el mínimo) una función objetivo, la cual depende de ciertas variables. Cada una de estas variables tiene un determinado dominio y está sujeta a restricciones. La solución está asociada a una sucesión de decisiones. En cada etapa se toma una decisión que no tiene vuelta atrás, y que involucra elegir el elemento más promisorio (la mejor opción de las disponibles, el candidato que parece ofrecer más posibilidades para mejorar la función objetivo) de un conjunto y se analizan ciertas restricciones asociadas a cada variable de la función objetivo para ver si se verifican mejoras en ella.

La estrategia greedy, muy interesante en sí, no sirve para cualquier problema: se debe tratar de un problema de optimización; pero aún en este caso, no asegura que se llegue al óptimo global (podría llevarnos a un óptimo local) ni tampoco, en algunos casos, llegar a una solución factible: el problema en cuestión debe verificar ciertas condiciones para que quede garantizado que la estrategia greedy funcione (estas condiciones se verifican para el problemas de los caminos mínimos con origen dado).

El problema a tratar tiene, entonces este enunciado:

Dado un grafo dirigido $G = (V, A)$, con aristas ponderadas, y considerando un vértice como origen, determinar los costos de los camino mínimos desde el vértice considerado origen a todos los otros vértices de V .

Resolución de Dijkstra utilizando una matriz de pesos y una tabla de vértices visitados:

Consideramos el grafo orientado G con aristas ponderadas no negativas. Vamos a considerarlos vértices etiquetados con valores $1..n$ y también que el vértice de partida es el 1.

La función a minimizar es la de costes desde el origen a cada uno de los restantes vértices.

En un vector D vamos a almacenar los costes actualizados a medida que avanza la ejecución del algoritmo.

Nuestro conjunto de candidatos es el de los vértices del grafo no visitados. Desde el comienzo en ese conjunto no estará el vértice de inicio, ya que de ahí partimos.

En cada etapa del algoritmo seleccionamos un vértice cuyo coste desde el origen sea mínimo. Este coste inicialmente es el que indican las aristas que parten del vértice origen, pero se va actualizando porque podemos bajarlo “pasando” por otro vértice de coste menor antes.

El algoritmo termina cuando el conjunto de candidatos queda sin elementos.

// $G = (V, A)$ es el grafo dirigido con aristas no negativas

//M es la matriz de pesos

//S es el conjunto de vértices visitados a lo largo del procesamiento.

// D es un vector de n-1 elementos para calcular los valores de los respectivos caminos mínimos

Caminos_Minimos_Dijkstra

{

// comenzamos con el vértice origen en el conjunto de vértices procesados

inicializar el conjunto S en $\{1\}$

Para cada vértice i desde 2 hasta N hacer

{

//si hay arista $(1, i)$ almacenamos su peso en la posición correspondiente de D

// o bien un peso máximo si no hay arista

$D[i] = M[1][i]$

}

//mientras haya candidatos que considerar, elegimos el más promisorio y analizamos condiciones

Mientras (conjunto de vértices no visitados $V-S$ no esté vacío)

{

Elegir vértice w perteneciente a $V - S$ tal que $D[w]$ sea mínimo

//Agregamos w al conjunto de vértices que fueron visitados

$S = S \cup \{w\}$

para todo vértice v adyacente a w

{

//Se registran las eventuales mejoras en aquellos v

$D[v] = \min(D[v], D[w] + M[w, v])$

}

}

}

Consideraciones sobre el coste temporal:

Si se utiliza una matriz de adyacencia y una tabla para almacenar los vértices y sus marcas de visitados, resulta:

$O(v)$ la carga inicial de D y la marca de no visitado para los n-1 vértices candidatos.

El ciclo que se ejecuta mientras haya candidatos, es obviamente $O(v)$.

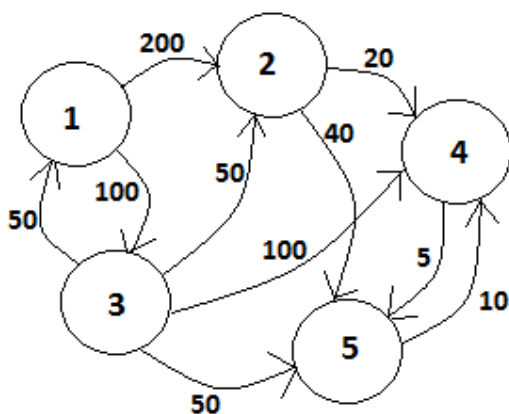
La función de elección de un vértice de coste mínimo no visitado implica recorrer tanto la tabla de visitados como el vector D, es $O(v)$.

El análisis de cada adyacente al nodo elegido, por estar utilizándose una matriz de pesos, implica recorrer la fila correspondiente al vértice elegido, es decir $O(v)$.

Aplicando las conocidas reglas de la suma y del producto, se llega a que el algoritmo es $O(v^2)$.

Ejemplo:

Consideraremos un grafo y los valores asociados a los arcos los presentaremos en una tabla; pero hay que recordar que la implementación del grafo podría tener estos datos en una matriz o en listas de adyacencia. Usaremos la versión del algoritmo desarrollada antes. Por ejemplo:



La tabla o matriz de pesos M es:

	1	2	3	4	5
1	0	200	100	∞	∞
2	∞	0	∞	20	40
3	50	50	0	100	50
4	∞	∞	∞	0	5
5	∞	∞	∞	10	0

El vector D donde registramos los sucesivos avances en el algoritmo, inicialmente es:

2	3	4	5
200	100	∞	∞

Hemos indicado en una fila las etiquetas de los nodos y debajo, sus costos. Estos valores iniciales los tomamos de la fila correspondiente al origen, en la table T. Los valores ∞ indican que no hay arista de enlace entre esos vértices.

El conjunto de candidatos V-S es {2, 3, 4, 5}. El vértice “más promisorio” de los que no fueron visitados es el de coste más bajo, en este caso el 3.

Seleccionamos 3, lo sacamos del conjunto de candidatos, y analizamos cuales son los adyacentes a 3 para ver si alguno puede tener mejora.

3 tiene como adyacentes a 1, 2, 4 y 5.

El coste de alcanzar 2 desde 3 es 50, por tanto, podemos llegar desde 1 hasta 2 pasando por 3 a coste 150 (100 es el coste de la arista (1,3) y 50 el coste de la arista (3, 2)). Este valor mejora el 200 de la table, por lo que lo mejoramos.

El coste de alcanzar 4 desde 3 es 100, entonces podemos llegar desde 1 hasta 4 a coste 20 (100 coste de arista (1, 3) más 100 de coste de arista (3,4)). Antes el coste era ∞ porque no había camino, por tanto mejoramos la situación.

El coste de alcanzar 5 desde 3 es 50, así que podemos llegar desde 1 hasta 5 pasando por 3 a coste 150 (100 es el coste de la arista (1,3) y 50 el coste de la arista (3, 5)). Esto mejora el ∞ anterior, mejoramos el valor.

Entonces el vector de costes luego de esta primera etapa queda:

2	3	4	5
150	100	200	150

De los vertices no visitados, (el conjunto es ahora {2, 4, 5}) hay dos de coste mínimo: 2 y 5. Podemos elegir cualquiera de ellos. Por ejemplo, el 2, que sacamos del conjunto de candidatos.

Con el 2, resulta lo siguiente: los vértices adyacentes son 4 y 5.

El costo de alcanzar 4 desde 2 es 20. Por lo que se puede alcanzar 4 desde 1 a coste 170, ya que 150 es el costo (1, 2) y 20 es el costo (2,4).

El costo de alcanzar 5 desde 2 es 40. Así que se alcanzaría 5 desde 1 pasando por 2 a coste 190, porque es 150 el costo de (1,2,) más 40 de costo de (2, 5). Pero el costo previo era 150, el nuevo valor (190) no lo mejora, por lo tanto no lo registramos, no es una mejora.

Entonces el vector de costes queda:

2	3	4	5
150	100	200	150

Ahora, de los vertices no visitados ({4, 5}) el más promisorio es el 5.

5 tiene como adyacente a 4. El coste de la arista (5, 4) es 10. Entonces puede alcanzarse el vértice 4 desde 1 a un coste más bajo que el que teníamos. Ahora, tenemos 150 para llegar a 5 desde 1 más 10 para alcanzar 4 desde 5.

Registramos la mejora

2	3	4	5
150	100	200	150

Nos ha quedado solo el 4 en el conjunto de candidatos. Lo retiramos y hacemos los análisis correspondientes, pero observamos que no se produce ninguna mejora.

Los costos mínimos hasta 2, 3, 4 y 5 son 150, 100, 200 y 150 respectivamente.

Variante más eficiente del algoritmo de Dijkstra, con listas de adyacencia y cola con prioridades:

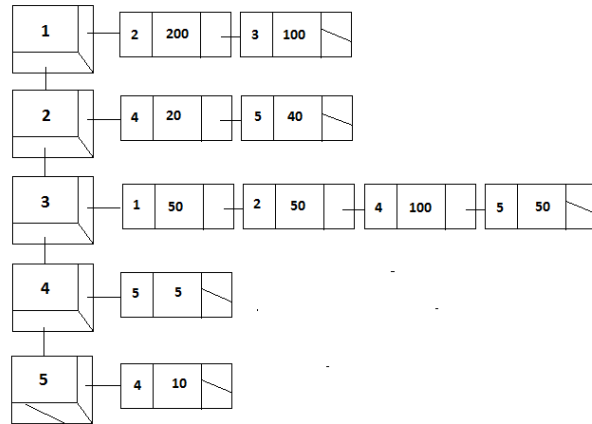
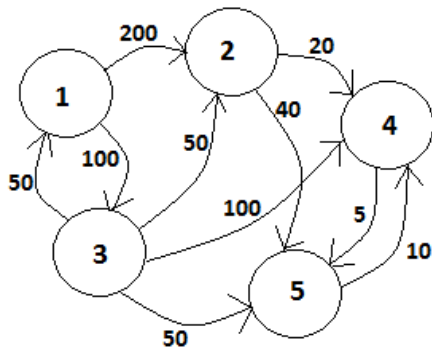
En el algoritmo que hemos descripto hay dos circunstancias que influyen mucho en la eficiencia: uno de ellos es la función de elección del más promisorio. Si tenemos los vértices almacenados en un array “común” con sus costes, hay que recorrer el array en cada etapa para establecer cual es el no visitado de coste mínimo. Eso implica un coste $O(n)$ en cada etapa. Puede disminuirse usando por ejemplo una cola con prioridad, que puede almacenarse en un heap de mínimo implementado en array. En esta cola la prioridad está dada por el coste de alcanzar el vértice desde el origen. Esto disminuiría el coste de la elección a $O(\log n)$, puesto que la elección del más promisorio corresponde a la baja de la raíz del heap, que puede hacerse en tiempo logarítmico.

Ahora bien, hay que considerar también que cada vez que se registre una mejora en el coste de alcanzar un vértice, si el vértice en cuestión está en el heap, se debe registrar la mejora, lo cual se refleja en un movimiento del nodo del heap, que se va mover en dirección a la raíz (lo reubicaremos comparándolo con el padre e intercambiando si es necesario, y luego comparando con el nuevo padre y así sucesivamente, como hacemos en un alta en el heap, pero en este caso con un nodo que ya estaba).

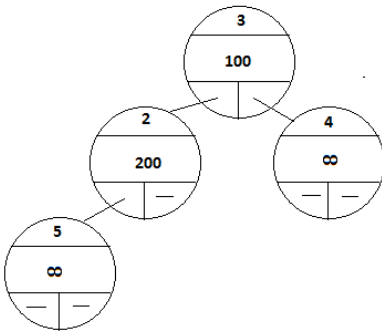
Además podemos implementar el grafo en listas de adyacencia, de modo que cuando determinemos quienes son los adyacentes de un vértice, en lugar de revisar $n-1$ celdas (que pueden no ser de adyacentes), sólo recorramos la lista de adyacentes.

Inicialmente se construye el heap de vértices con sus costes iniciales (esto, como hemos visto antes, puede hacerse a coste $O(n)$) y se inicializa el vector de costes con los valores de la lista de adyacentes del vértice origen.

Mientras el heap con candidatos no se vacíe, se elige al mejor (el de la raíz) y se restaura el heap. Para cada adyacente al elegido (lo cual se realiza usando la correspondiente lista de adyacentes) se analiza si hay una mejora, de haberla se registra y si el vértice mejorado está en el heap, se modifica su prioridad en el heap y se restaura nuevamente la estructura.



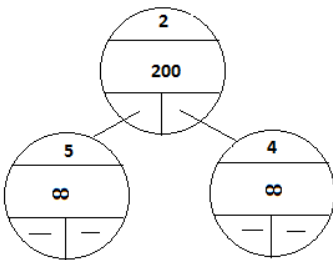
Por otro lado, el heap con los candidatos y sus costos inicialmente es:



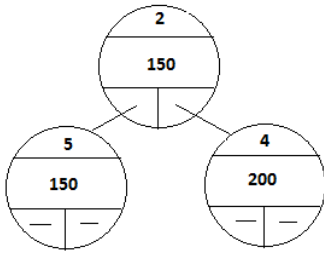
La tabla de costes inicial es

2	3	4	5
200	100	∞	∞

Eliminamos la raíz del heap y lo recomponemos; quedará así:



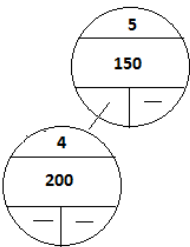
El 3, valor elegido, permite mejorar el valor de 2, a 150, luego 4 cambia a 200 y 5 a 150. Estas modificaciones se reflejan en el heap, que quedará así:



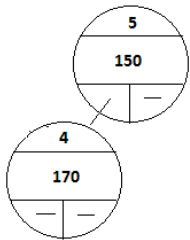
Y el vector de costos quedará:

2	3	4	5
150	100	200	150

Nuevamente eliminamos la raíz del heap y lo recomponemos; queda:



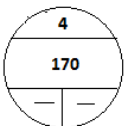
Recorremos la lista de adyacentes a 2, se mejora 4 y esto se refleja en el heap, que queda:



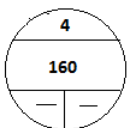
El array de costos quedará así:

2	3	4	5
150	100	200	150

Eliminamos nuevamente la raíz del heap, queda



Con 5 el coste de llegar a 4 baja, quedando el heap así:



Finalmente eliminamos le ultimo nodo del heap, pero no se obtiene ninguna mejora.
El array, como habíamos visto queda finalmente así:

2	3	4	5
150	100	200	150

Análisis del coste temporal:

El armado inicial del heap se puede hacer cargando los costos de los vértices ($2..n$) en un array y construyendo un heap sobre él, lo cual puede lograrse en $O(v)$.

Mientras el heap no se haya vaciado, se elimina la raíz y se restaura el heap. En total esto se hará v veces, a un costo total $O(v * \log v)$.

Para cada adyacente al elegido puede tener que modificarse, en el peor caso, las prioridades de todos los vértices que estén en el heap, el cual deberá ser restaurado. Esto tiene un coste de $O(a * \log v)$.

Ahora bien, hay que tener en cuenta que la operación de determinar si el vértice cuyo costo se modifica está en el heap y en qué posición del heap se encuentra, tiene un coste asociado. Si este coste fuera $O(1)$, entonces, el costo total es: $O(v) + O(v * \log v) + O(a * \log v)$, es decir $O((a+v) \log v)$.

Problema de los caminos más cortos entre todos los pares de vértices.

Se trata de determinar los caminos mínimos que unen cada vértice del grafo con todos los otros. Si bien lo más común es que se aplique a grafos que no tengan aristas negativas, la restricción necesaria es que no existan ciclos con costos negativos.

Este problema suele resolverse aplicando la estrategia de “Programación Dinámica”.

La programación dinámica suele aplicarse también a problemas de optimización. En estos problemas se realiza una división de problemas en otros problemas menores, pero, a diferencia de lo que ocurría en “Divide y Vencerás”, estos problemas resultado de la división no son independientes entre sí, sino que tienen subproblemas en común, hay un ‘solapamiento’ de problemas. La técnica consiste en resolver y almacenar las soluciones de las zonas solapadas para no volver a realizar los mismos cálculos. En general en esta estrategia se llega a la solución realizando comparaciones y actualizaciones en datos que han sido tabulados (las soluciones de los subproblemas).

Algoritmo de Floyd

Este algoritmo desarrollado en 1962 (<https://dl.acm.org/citation.cfm?doid=367766.368168>) se basa en una técnica de programación dinámica que almacena en cada iteración el mejor camino entre el que pasa por el nodo intermedio k y el que va directamente del nodo i al nodo j .

Para determinar cuál es el mejor camino considera mínimo ($D[v]$, $D[u] + M(u,v)$).

Inicialmente se carga una matriz de $N \times N$ con los pesos correspondientes a las aristas.

El método lleva a cabo N iteraciones del proceso en cada una de las cuales se evalúa la eventual mejora de los tiempos por la consideración del vértice i (que variara entre 1 y N a lo largo del proceso) como intermedio.

Algoritmo de Floyd

// N es el número de vértices

{

Inicializar A con ceros en la diagonal principal y los pesos indicados por el grafo para cada uno de los restantes elementos. Si una arista no existe, se coloca peso infinito.

Para k desde 1 hasta N

 Para i desde 1 hasta N

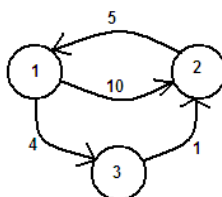
 Para j desde 1 hasta N

$A[i][j] = \text{mínimo} (A[i][j], A[i][k] + A[k][j])$

}

El coste temporal de este algoritmo es, obviamente $O(N^3)$.

Ejemplo: Consideremos el siguiente grafo para aplicar el algoritmo de Floyd



Matriz A

	1	2	3
1	0	10	4
2	5	0	∞
3	∞	1	0

Pasando por el vértice 1, A queda así

	1	2	3
1	0	10	4
2	5	0	9
3	∞	1	0

Pasando por el vértice 2, A queda así

	1	2	3
1	0	10	4
2	5	0	9
3	6	1	0

Pasando por el vértice 3, A queda así

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

Cerradura transitiva:

Conjunto de pares de vértices de un grafo orientado cuyas componentes son vértices ligados por algún camino (no importa la longitud del mismo). Algoritmo de Warshall (1962). Algoritmo de Warshall:

{

 Inicializar la matriz de adyacencia del grafo

```
Para k desde 1 hasta n
{
  Para i desde 1 hasta n
  {
    Para j desde 1 hasta n
    { A[i][j] = Maximo ( A[i][j] , A[i][k] * A[k][j]);}
  }
}
```

Ejercicios propuestos:

1. Dado un grafo dirigido y acíclico con pesos no negativos en las aristas, se quiere determinar la distancia máxima desde un vértice origen a cada uno de los otros. ¿Se puede diseñar un algoritmo similar al de Dijkstra eligiendo al candidato de coste mayor en cada etapa?
2. Modificar el algoritmo de Floyd para calcular el número de caminos con distancia mínima que hay entre cada par de nodos. ¿Cómo queda el coste del algoritmo?
3. Mostrar mediante un contraejemplo que el algoritmo de Dijkstra no sirve si el grafo tiene alguna arista negativa.

Algunos problemas sobre Grafos no dirigidos.

Como ya se ha definido, un árbol libre es un grafo no dirigido conexo sin ciclos.

Se verifica que

- En todo árbol libre con N vértices ($N > 1$), el árbol contiene $N-1$ aristas.
- Si se agrega una arista a un árbol libre, aparece un ciclo.
- Si u y v son dos vertices distintos de un árbol, entonces hay un solo camino que los une.

Un grafo con peso o ponderación en las aristas es un grafo $G = (A, V)$, con un costo $C(v, w)$, asociado a cada arista (v, w) perteneciente a A .

La representación por matriz de adyacencia la haremos colocando el peso de la arista en la celda de la misma.

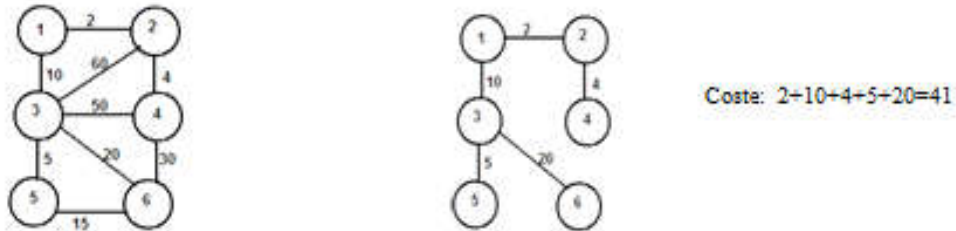
Para la representación por lista de adyacencia, agregaremos un atributo peso a cada nodo de la lista de adyacentes.

Definimos como árbol abarcador de un grafo $G = (V, A)$ no dirigido y conexo a un árbol libre que conecta todos los vértices de V .

Arbol Abarcador de Coste Mínimo:

Dado un grafo $G=(V,A)$ no dirigido con aristas ponderadas (es decir, que cada arista (v, w) de A tiene un costo asociado, $C(v, w)$), se llama árbol abarcador de coste mínimo al árbol abarcador para G que verifica que la sumatoria de los pesos de las aristas es mínima. (El árbol abarcador de costo mínimo, puede no ser único).

Ejemplo: para el grafo de la izquierda, un árbol abarcador de coste mínimo es el de la derecha.



Algoritmos que permiten obtener el arbol de expansion de coste minimo para un grafo dado

Existen varios algoritmos que permiten obtener el árbol de expansión de coste mínimo.

Se detallaran dos de ellos: el algoritmo de Prim y el de Kruskal.

En ambos casos se aplica la estrategia 'greedy' o voraz para alcanzar el óptimo.

Algoritmo de Prim

Desarrollado por Prim en 1957 sobre un algoritmo previo de Jarnik de 1930.

(<https://archive.org/details/bstj36-6-1389>)

Usa una estrategia greedy.

Las aristas deben ser no negativas.

Conjunto de candidatos: conjunto de vértices del grafo aún no incorporados al árbol (cualquiera de ellos se elige como inicial)

Función de selección: elección del vértice que aún no está en el árbol para el cual existe alguna arista de peso mínimo que lo enlaza con algún vértice del árbol.

Función a optimizar: sumatoria de pesos de aristas del árbol

Criterio de finalización: todos los vértices están en el árbol.

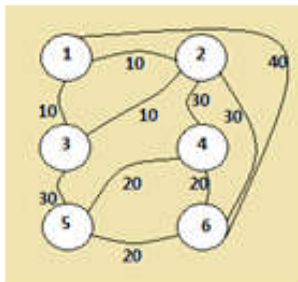
Esquema de Prim:

Llamando G al grafo y A al conjunto de aristas del árbol abarcador de coste mínimo, es

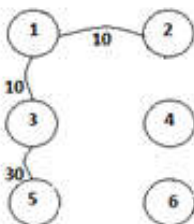
```
{
  A se inicializa en vacío
  U = {1} //se almacena vértice inicial
```

Mientras U distinto de V hacer:

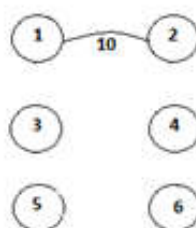
```
{
  Determinar la arista de costo mínimo,  $(u,v)$  tal que  $u \in U$  y  $v \in (V-U)$ 
  Agregar a  $A$  la arista  $(u, v)$ 
  Agregar a  $U$  el vértice  $v$ 
}
```



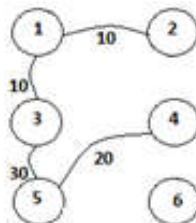
Grafo inicial



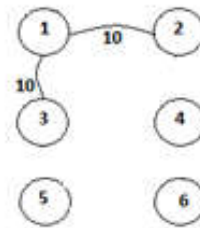
Paso 3



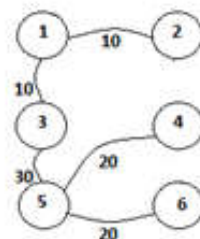
Paso 1



Paso 4

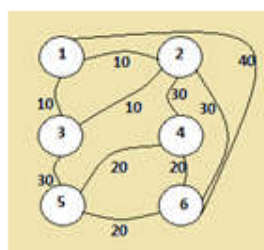


Paso 2

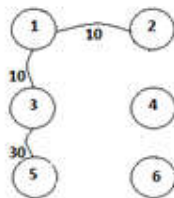


Paso 5: árbol abarcador de coste mínimo.
Coste: 90

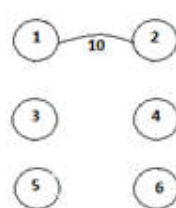
Algoritmo de Kruskal



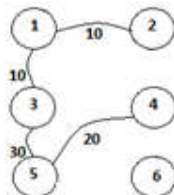
Grafo inicial



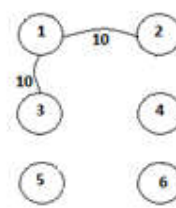
Paso 3



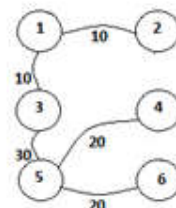
Paso 1



Paso 4



Paso 2



Paso 5: árbol abarcador de coste mínimo.
Coste: 90

Desarrollado por Kruskal en 1956.

Usa una estrategia greedy. Las aristas deben ser no negativas.

Conjunto de candidatos: conjunto de aristas del grafo.

Función de selección: elección de la arista de coste mínimo.

Restricción: la arista no debe formar ciclo.

Función a optimizar: sumatoria de pesos de aristas del árbol

Criterio de finalización: el árbol tiene $n-1$ aristas.

Esquema del algoritmo:

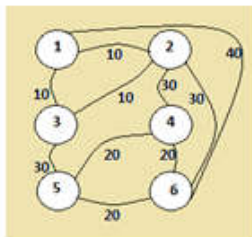
//G.(V, A) grafo no dirigido con aristas no negativas

//Consideraremos los vértices etiquetados como 1..n. Se comenzará incorporando al vértice 1 (arbitrariamente)

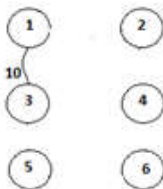
//E: conjunto de aristas del árbol abarcador; U:conjunto de vértices del árbol

```
{
  Ordenar aristas de forma creciente
  Inicializar E en vacío
  Inicializar U en vacío
  Mientras #E < n-1
  {
    Elegir arista (u,v) de peso mínimo que no forme ciclo
    E = E U {(u,v)}
    Incorporar al conjunto U el vértice u , o el V, según corresponda
  }
}
```

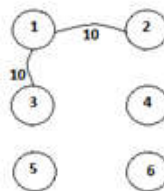
Ejemplo: para el grafo ya considerado,



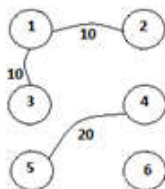
Grafo inicial



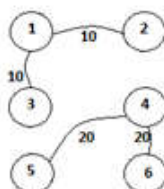
Paso 1



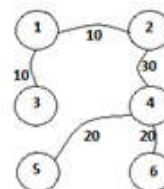
Paso 2



Paso 3
La arista de coste 10
entre 2 y 3 forma
ciclo, se descarta



Paso 4
La arista de coste 20
entre 5 y 6 se descarta



Paso 5: árbol abarcador
de coste mínimo.
Coste: 90

¿Cómo chequear si cada nueva arista considerada (u,v) forma ciclo?

u y **v** en la misma componente => incorporar (u, v) genera ciclo.

Se comienza con una componente para cada vértice. Las componentes se unen al añadir una arista.

Función que retorna componente de un vértice v: Find(v)

Función que une la componente de u y la componente de v: Union(u, v).

(Se demuestra que ejecutar k uniones con n elementos tiene un coste temporal $k \log(n)$, ya que en k uniones hay $2*k$ elementos involucrados)

Kruskal //versión con Union y Find para un grafo de n vértices

Inicializar A conjunto de aristas del árbol abarcador en vacío

Ordenar S conjunto de aristas del grafo de forma creciente

Para cada vértice v

{MakeSet (v);}

Mientras (#A < n-1)

{

Elegir arista (u, v) de peso mínimo y retirarla de S

si (Find (u) != Find (v))

{

Union (Set(u) , Set(v));

A = A U {(u,v)};

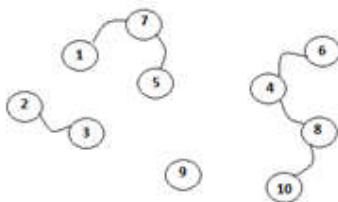
}

}

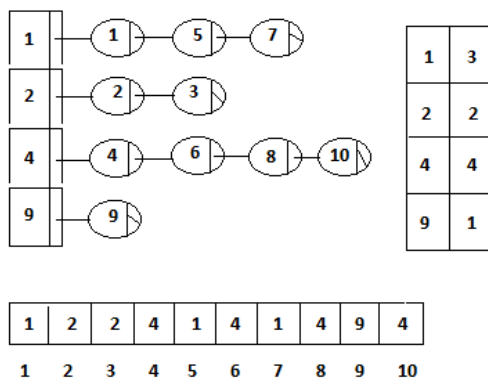
Posibles implementaciones para Union y Find:

Las diferentes formas de implementar las funciones Union y Find influyen en el costo del algoritmo de Kruskal.

Ejemplo, para el siguiente grafo,



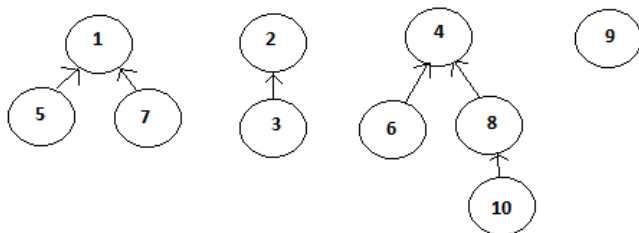
Puede plantearse una implementación de este tipo,



En esta implementación, para las componentes hay un array de listas. El menor valor es representante de la lista. Para unir se elige la lista menor que se incorpora a la otra. Para determinar los tamaños, hay una tabla con los mismos (a la derecha)

El array inferior permite implementar Find con $O(1)$.

Otra posibilidad es



En esta implementación, para las componentes hay estructuras arborescentes “invertidas”. En cada una, cada nodo, excepto uno, apunta a su padre. El que está en el lugar de la raíz es el representante de la componente.

Ejercicios propuestos:

1. Analizar el coste del algoritmo de Kruskal con las estructuras sugeridas para implementar Union y Find.
2. ¿Cuál de los dos algoritmos (Prim y Kruskal) conviene más si la densidad de aristas es alta?
3. Analizar la validez, y en ese caso el coste del siguiente algoritmo para determinar las componentes conexas de un grafo no dirigido:

Para cada vértice v del grafo {MakeSet (v);}

Para cada arista de (x,y) perteneciente a A hacer

```
{  
  Si (Find( $x$ ) == Find ( $y$ ))  
    { Union ( $x$ ,  $y$ ); }  
}
```

¿Cuántas veces se ejecuta Union?

¿Y Find?