Programación II

Tecnicatura Universitaria en Inteligencia Artificial

2023

Teoría de grafos

Índice

1. Grafos simples

1.1. Ejemplo motivador

La Figura 1 muestra un recorte de un mapa de la provincia de Santa Fe, particularmente, de Rosario y alrededores. Dicho mapa señala algunas ciudades del sur de la provincia, junto con algunas de las rutas más importantes que las conectan.

Figura 1: Mapa de Rosario y sus alrededores

Supongamos que, como parte de un proyecto de mejoramiento del estado de los caminos, la provincia de Santa Fe desea recorrerlos para determinar en dónde sería conveniente invertir en luminaria y señalética. Una cuestión a resolver es cómo armar el recorrido que se hará, de modo que sea lo más económico posible (en tiempo y nafta). Siendo que el recorrido comienza en Rosario (porque el equipo de trabajo vive allí), lo ideal sería recorrer cada camino exactamente una vez y regresar de nuevo al punto de partida. Lo que no sabe el encargado del proyecto (porque desgraciadamente nunca estudió teoría de grafos) es que, en este caso, hacer eso es imposible. Analicemos esta situación.

Este problema puede *modelarse* como un **grafo**. Los grafos son diagramas con dos componentes fundamentales: puntos y líneas (que conectan dichos puntos). Los puntos se llaman *vértices* o *nodos*, mientras que las líneas se llaman *aristas*. En la Figura 2 se muestra el mismo mapa representado como un grafo. Los nodos representan a las ciudades, mientras que las aristas indican la existencia de una ruta entre las ciudades que une. Notemos que las rutas que no unían ciudades indicadas no aparecen en el grafo (porque no son pertinentes al problema), y que, tanto Granadero Baigorria como Villa Gobernador Galvez, se representan en el mismo punto que Rosario.

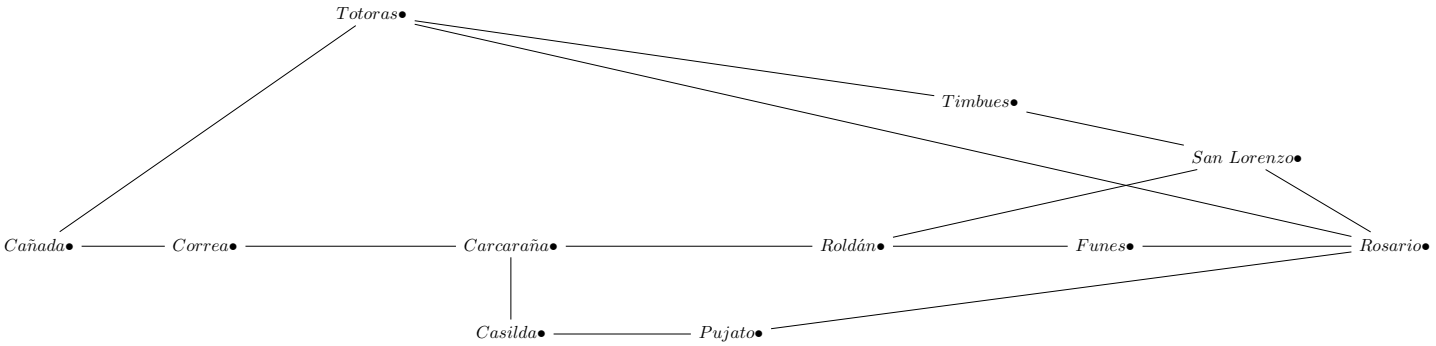


Figura 2: Rosario y sus alrededores representado como grafo

Lo primero que debe notarse es que en un grafo solo se está representando la existencia de ciertos nodos, y las conexiones entre ellos. Existen otros aspectos del grafo que podríamos observar en la figura de arriba. Por ejemplo, la posición relativa de los nodos (Roldán está a la izquierda de Funes), o la longitud de las distintas aristas (Pujato está más cerca de Casilda que de Rosario). Sin embargo, estos aspectos no aportan información significativa al modelo (¡ni son pertinentes para nuestro problema, por eso un grafo es útil para resolverlo!). Recordemos que lo que queremos encontrar es un recorrido que comience en Rosario, recorra cada arista una única vez, y retorne al punto de partida. Así, por ejemplo, el grafo de la Figura 3 es igualmente válido para representar el problema en cuestión (aunque Roldán ya no esté a la izquierda de Funes, o Pujato ya no no esté más cerca de Casilda que de Rosario).

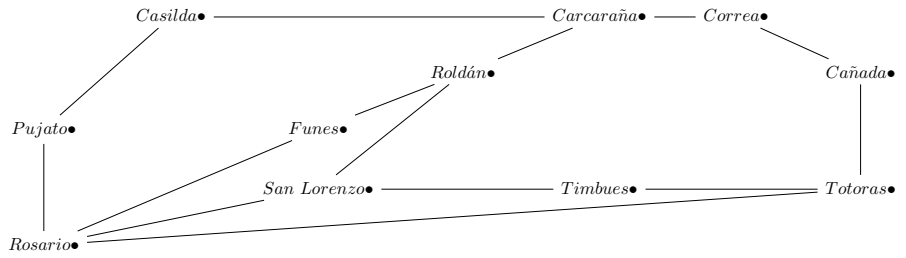


Figura 3: Rosario y sus alrededores representado como grafo (alternativa)

Gracias a la teoría de grafos, es posible demostrar que tal recorrido que estamos buscando es imposible de lograr en una situación como la que tenemos. Volveremos a esto en futuras secciones.

A lo largo de este apunte veremos definiciones y conceptos fundamentales de la teoría de grafos, algunos algoritmos famosos, y una librería de `python` muy útil para trabajar con ellos.

1.2. Conceptos básicos sobre grafos

Definición. Un **grafo** (también llamado grafo no dirigido o grafo simple) G consiste en un conjunto de vértices (o nodos) y un conjunto E de aristas (o arcos) tal que cada arista $e \in E$ se asocia con un **par no ordenado** de vértices. Si la arista e está asociada con los vértices v y w, se escribe $e = (v, w)$ o $e = (w, v)$. Es importante notar que ambas son exactamente la misma arista.

El grafo de la Figura 3 es un grafo simple: *Casilda* y *Pujato* son ejemplos de nodos, y el par no ordenado $(Casilda, Pujato)$ es una arista del grafo.

Definición Se dice que una arista $e = (v, w)$ en una grafo es *incidente* sobre v y w. A su vez, se dice que los vértices v y w son incidentes sobre e y que son **adyacentes** (o *vecinos*) entre ellos.

Por ejemplo, $Funes$ y *Roldán* son vértices vecinos, mientras que $Pujato$ y $Totoras$ no lo son.

Si G es un grafo con vértices V y aristas E, se escribe $G = (V, E)$. Usualmente, se pide que los conjuntos E y V sean finitos, y que V sea no vacío.

El grafo de la Figura 3 se puede definir como $G = (V, E)$, donde

- $V = \{Casilda, Caracaraña, Correa, Roldán, Cañada, Pujato, Funes, San\ Lorenzo, Timbues, Totoras, Rosario\}$
- $E = \{(Casilda, Caracaraña), (Carcaraña, Correa), (Casilda, Pujato), (Carcaraña, Roldán), (Correa, Cañada), (Cañada, Totoras), (Roldán, Funes), (Pujato, Rosario), (Funes, Rosario), (Roldán, San\ Lorenzo), (San\ Lorenzo, Timbues), (Timbues, Totoras), (Rosario, San\ Lorenzo), (Rosario, Totoras)\}$

Definición. Sea un grafo $G = (V, E)$, un **camino** en dicho grafo es una sucesión de vértices $v_0, v_1, ..., v_n$, y aristas $l_0, l_1, ..., l_{n-1}$, tales que $l_i = (v_i, v_{i+1})$ para todo $i \in \{0, 1, ...n\}$. La **longitud** del camino se determina por la cantidad de aristas que recorre, en este caso, n.

En la Figura 4 se muestra el camino formado por la sucesión de vértices *Rosario*, *Funes*, *Roldán*, *San Lorenzo*, *Timbues*, *Totoras*, y las aristas $(Rosario, Funes)$, $(Funes, Roldán)$, $(Roldán, San\ Lorenzo)$, $(San\ Lorenzo, Timbues)$, $(Timbues, Totoras)$. La longitud de dicho camino es 5.

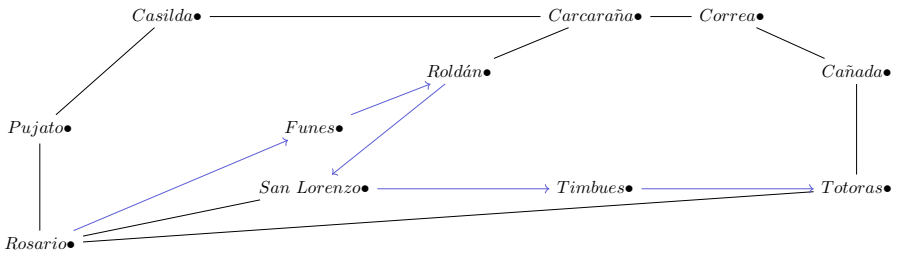


Figura 4: Camino

Observe que la definición de camino permite tener aristas repetidas. Llamaremos **camino simple** a un camino donde no hay aristas repetidas. Además, llamamos **ciclo** a un camino que empieza y termina en el mismo vértice. Un **ciclo simple** es un ciclo que no repite aristas.

1.3. Retomando el ejemplo motivador

Ahora que sabemos conceptos básicos sobre teoría de grafos, podemos escribir el problema del ejemplo introductorio formalmente como: ¿Existe un camino desde Rosario hacia Rosario que pase por cada arista exactamente una vez?

Es posible demostrar que, para desgracia del encargado el proyecto, no es posible construir un camino tal. Para convencernos de esto intuitivamente, supongamos que existe tal trayectoria, y consideremos el vértice *Funes*. Cada vez que se llega a *Funes* por alguna arista, se debe salir de *Funes* por una arista diferente. Más aún, cada arista que toca *Funes* se debe usar. Entonces las aristas en *Funes* ocurren en pares. Se concluye que un número par de aristas debe tocar *Funes*. Como tres aristas tocan a *Funes*, se tiene una contradicción. Por lo tanto, no existe un camino del vértice *Rosario* al vértice *Rosario* en el grafo que pase por cada arista una única vez. El mismo razonamiento se puede aplicar a un grafo arbitrario.

2. Grafos dirigidos

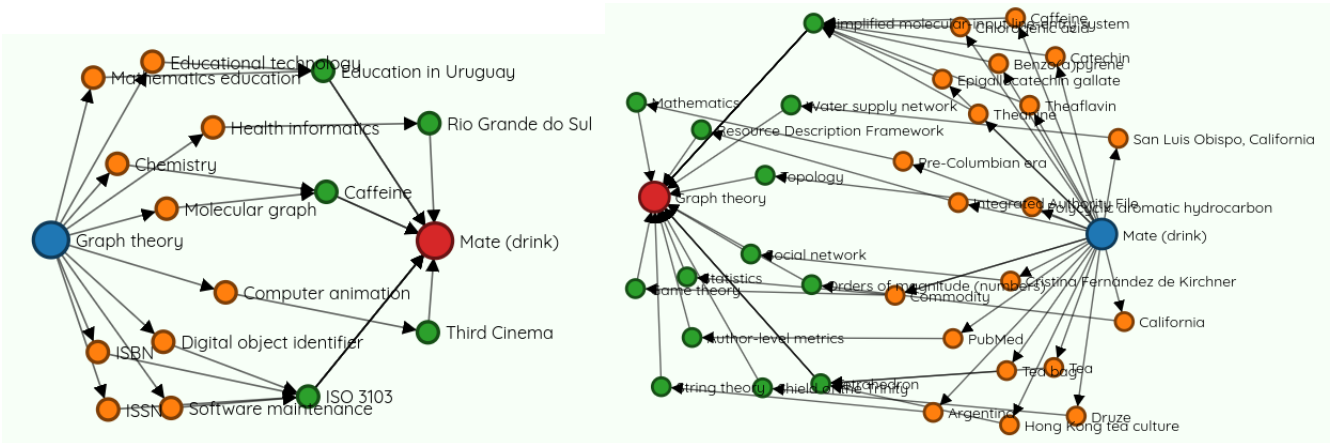
2.1. Ejemplo motivador

Wikipedia es una enciclopedia libre, políglota y editada de manera colaborativa. Fue creada el 15 de enero de 2001, y es la mayor y más popular obra de consulta en Internet. Una de las características que hace tan fácil la navegación a través de sus artículos son sus enlaces (o *links*) internos. Los enlaces internos son hipervínculos a otro recurso (por ejemplo, un artículo) dentro del mismo dominio. Además de permitir a los usuarios navegar de un tema a otro de manera rápida y sencilla, los enlaces entre artículos conectan conceptos, lo que permite a los lectores profundizar en un tema específico o explorar temas relacionados.

El *wiki juego*[1], también conocido como la *carrera de Wikipedia*, es un juego en donde los participantes (uno o más) comienzan posicionados en el mismo artículo seleccionado arbitrariamente, y deben navegar hasta otro artículo pre-seleccionado. El objetivo es llegar al artículo destino en la menor cantidad de clicks. ¿Nos suena a algo?

El wiki juego puede modelarse como la búsqueda de un camino de longitud mínima entre dos vértices determinados de un grafo, en donde los vértices representan artículos y, las aristas, un enlace interno entre ellos. Sin embargo, hay algo novedoso a observar sobre los enlaces internos: son direccionales. A diferencia de las rutas de Santa Fe, que son bidireccionales (o doble-mano), los enlaces internos tienen un sentido definido. Así, por ejemplo, que el artículo Mate (infusión) de la Wikipedia en Español, tenga un enlace interno al artículo Estratificación social, no significa que necesariamente también pase a la inversa.

A modo de ejemplo, en la Figura 5 se muestra una representación de los enlaces internos de Wikipedia como grafos dirigidos. En particular, se muestran todos los caminos mínimos (de longitud 3) desde el artículo *Mate (drink)* hasta el artículo *Graph Theory*, y viceversa (ambos artículos de la Wikipedia en inglés).



(a) Caminos mínimos desde Graph Theory a Mate

(b) Caminos mínimos desde Mate a Graph Theory

Figura 5: Representación de enlaces internos de Wikipedia como grafos dirigidos.

2.2. Conceptualización formal

Definición. Un **grafo dirigido** G consiste en un conjunto V de vértices y un conjunto E de aristas tal que cada arista $e \in E$ se asocia con un **par ordenado** de vértices. Si la arista e está asociada

[1]Disponible para jugar en `https://www.thewikigame.com/`.

con los vértices u y w, se escribe $e = (v, w)$. Es importante considerar que en este contexto las aristas (v, w) y (w, v) representan cosas distintas. La Figura 6 muestra un grafo dirigido.

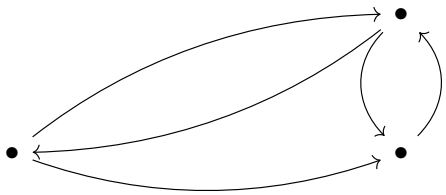
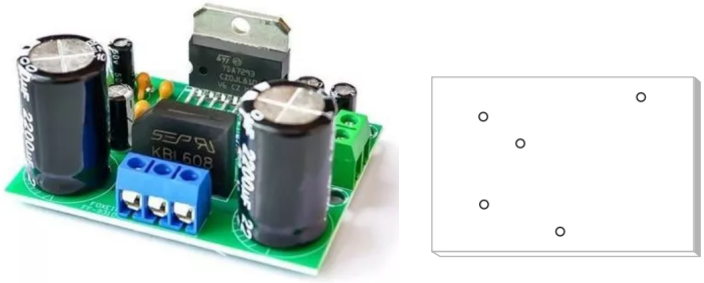


Figura 6: Grafo dirigido

3. Grafos ponderados

3.1. Ejemplo motivador

A menudo, en la manufactura, es necesario hacer agujeros en hojas de material específico para luego atornillar componentes a estas placas. Supongamos que una fábrica está diseñando un nuevo tipo de módulo amplificador. Como parte de este proceso, se deben diseñar las plaquetas electrónicas correspondientes. Dichas plaquetas necesitarán contar con ciertos agujeros, dispuestos en lugares específicos para, por ejemplo, la colación de las borneras requeridas. Los agujeros se perforan utilizando un taladro controlado por computadora. Para ahorrar tiempo y dinero, el taladro debe moverse tan rápido como sea posible. En la figura 7 se muestra un ejemplo de módulo amplificador junto con una representación simplificada del mismo.



(a) Ejemplo (b) Representación simplificada

Figura 7: Módulo amplificador

Para analizar el problema de minimización (o al menos, disminución) del tiempo utilizado en hacer todos los hoyos necesarios podemos utilizar un **grafo ponderado** como modelo. Los vértices del grafo corresponden a los agujeros. Cada par de vértices se conecta por una arista. En cada arista se escribe el tiempo que toma mover el taladro entre los hoyos correspondientes. En la Figura 8 se presenta dicho grafo.

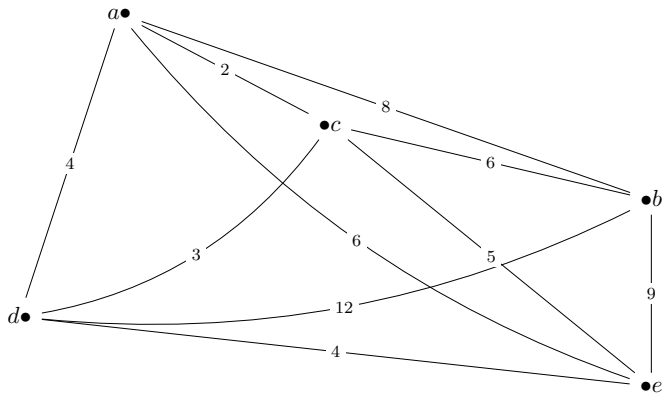


Figura 8: Módulo amplificador representado como grafo

Es interesante notar que este problema es muy distinto al analizado en la sección 1.1. Allí se buscaba encontrar un camino que comenzara y terminara en Rosario, y recorriera exactamente una vez cada una de las rutas de Santa Fe. Ahora no estamos interesados en un camino que recorra todas las aristas, sino en un camino que recorra todos los vértices exactamente una vez. Además, cada arista tiene un costo que hay que considerar, puesto que queremos encontrar el recorrido óptimo.

La tabla 1 muestra todas las rutas que comienzan en a y terminan en e, pasando por todos los nodos intermedios una vez.

Camino	Costo
a, b, c, d, e	21
a, b, d, c, e	28
a, c, b, d, e	24
a, c, d, b, e	26
a, d, b, c, e	27
a, d, c, b, e	22

Cuadro 1: Costos de recorrer todos los nodos comenzando en a y terminando en e

De todas las rutas presentadas, la que visita los vértices a, b, c, d, e, en ese orden, es la de longitud mínima. Por supuesto, un par diferente de vértices de inicio y fin podría producir una ruta aún más corta. Numerar todas las trayectorias posibles es un método muy ineficiente para encontrar el camino de longitud mínima que visita todos los vértices exactamente una vez. Por desgracia, nadie conoce un método que sea mucho más práctico para grafos arbitrarios. Este problema es una versión del *problema del agente viajero*. Se investiga actualmente para producir mejores aproximaciones a la respuesta, dado que la respuesta óptima es muy difícil de conseguir.

3.2. Conceptualización formal

Definición. Llamamos **grafo ponderado** o **grafo con pesos** a un grafo $G = (V, E)$ que cuenta con una función $w : E \to \mathbb{R}$, tal que $w(e)$ corresponde al peso de la arista e para todo $e \in E$. Es decir, un grafo ponderado es aquel para el que cada arista tiene asignado un peso. Notamos $w(i, j)$ al peso de la arista que va de i a j, si esta existe.

Definición. La **longitud de un camino** en un grafo con pesos es la suma de los pesos de las aristas que recorre dicho camino. Es importante notar que la definición es distinta a la dada para grafos simples.

4. Grafos como un TAD

Podemos pensar a los grafos como un TAD. Formalmente lo definimos de la siguiente forma:

Un Grafo es un conjunto de vértices (o nodos) y un conjunto de aristas (o arcos, o relaciones) entre ellos que admite la siguientes operaciones:

`__init__()` Que crea un grafo vacío.

`add_node(x)` Que agrega un nodo al grafo.

`remove_node(x)` Que remueve un nodo del grafo.

`add_edge(x,y)` Que agrega la arista entre los nodos x e y al grafo.

`remove_edge(x, y)` Que remueve la arista entre x e y del grafo.

`are_adjacent(x, y)` Devuelve True si x e y son adyacentes, False en caso contrario.

`is_node(x)` Devuelve True si x es un nodo del grafo, False si no.

`get_nodes()` Devuelve una lista de nodos del grafos.

`get_adjacent(x)` Devuelve una lista de todos los nodos adyacentes al nodo x.

En el caso de un grafo ponderado, definimos el TAD incorporando los cambios necesarios:

Un Grafo Ponderado es un conjunto de vértices (o nodos) y un conjunto de aristas (o arcos, o relaciones) entre ellos que admite la siguientes operaciones:

`__init__()` Que crea un grafo ponderado vacío.

`add_node(x)` Que agrega un nodo al grafo.

`remove_node(x)` Que remueve un nodo del grafo.

`add_edge(x,y, w)` Que agrega la arista entre los nodos x e y, con peso w, al grafo.

`remove_edge(x, y)` Que remueve la arista entre x e y del grafo.

`are_adjacent(x, y)` Devuelve el peso de la arista (x,y) si x e y son adyacentes, imprime un error en caso contrario.

`is_node(x)` Devuelve True si x es un nodo del grafo, False si no.

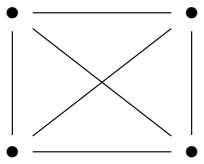
`get_nodes()` Devuelve una lista de nodos del grafos.

`get_adjacent(x)` Devuelve una lista de todos los nodos adyacentes al nodo x.

5. Grafos especiales

Algunos grafos aparecen con tal frecuencia que se les ha dado un nombre.

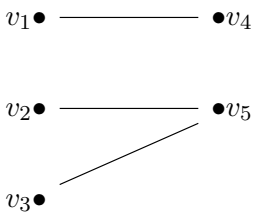
Definición El grafo completo de n vértices, notado como K_n, es el grafo simple con n vértices y una arista entre cada par de vértices distintos. A modo de ejemplo, se muestra el grafo K_4.



Definición Un grafo $G = (V, E)$ es un **bipartito** si existen subconjuntos V_1 y V_2 de V tales que:

1. $V_1 \cap V_2 = \emptyset$
2. $V_1 \cup V_2 = V$
3. Cada arista en E es incidente en un vértice V_1 y un vértice en V_2.

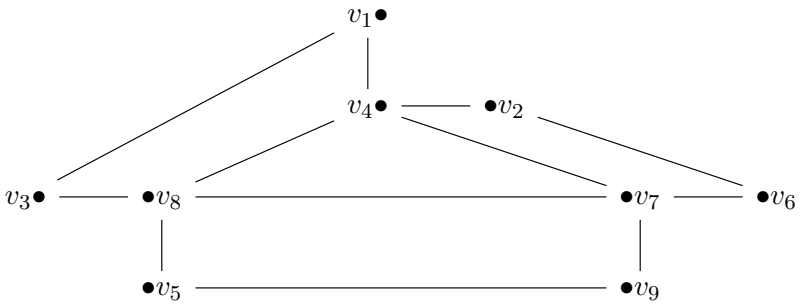
Por ejemplo, el grafo de la siguiente figura es bipartito con $V_1 = \{v_1, v_2, v_3\}$ y $V_2 = \{v_4, v_5\}$.



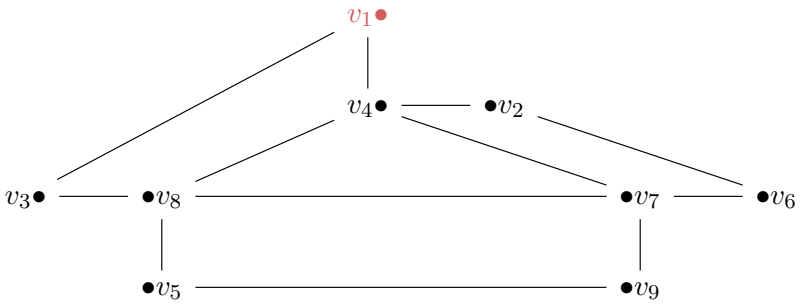
Ejemplo 1

Para identificar si un grafo es bipartito, conviene proceder por *coloreo* de dos colores de los vértices del grafo. La idea es pintar los vértices vecinos del grafo con colores progresivamente. Si al finalizar hemos pintado todo el grafo, concluimos que es bipartito, y los colores nos dan la asignación de V_1 y V_2 que necesitamos para que se cumpla la definición. Si, por el contrario, en algún momento encontramos un vértice que no podemos pintar de ninguno de los dos colores (porque tiene vecinos de ambos colores), hemos de concluir que el grafo no es bipartito.

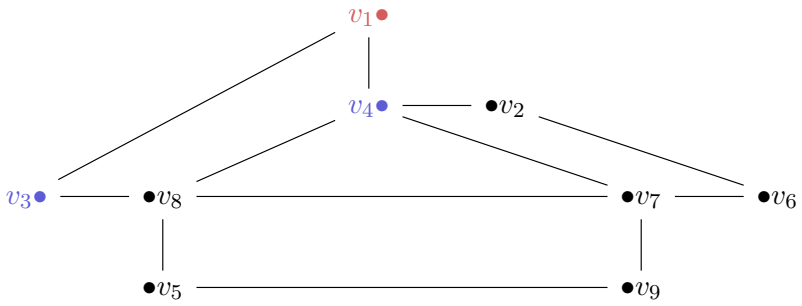
Utilicemos esta idea para decidir si el siguiente grafo es bipartito:



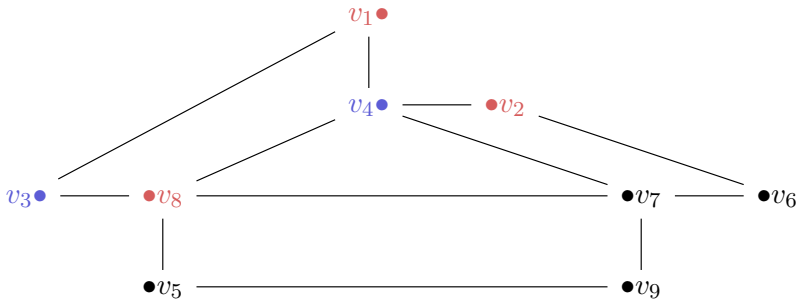
Comenzamos pintando cualquier vértice de un color.



En un segundo paso, pintamos todos sus vecinos de un color distinto:



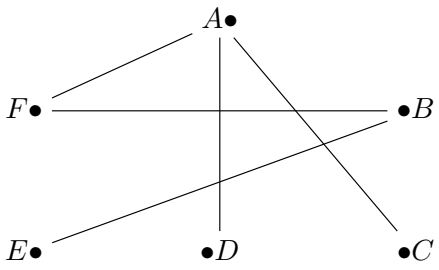
Continuando con este procedimiento, pintamos los vecinos de aquellos que acabamos de pintar del color original



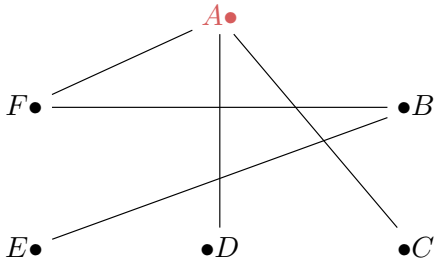
¿Que ocurre con el vértice v_7? Tiene vecinos de ambos colores, por lo tanto, no podemos pintarlo de ningun color. Concluimos que el grafo no es bipartito

Ejemplo 2

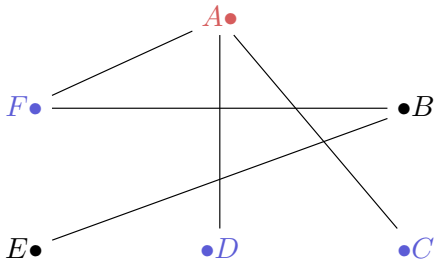
Consideremos ahora el siguiente grafo:



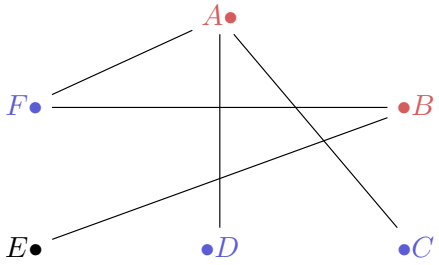
Nuevamente, para determinar si es bipartito, procedemos por coloración. Para empezar, pintamos el vértice A del primer color:



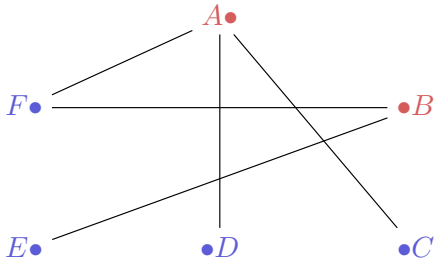
Luego, pintamos los vecinos de A del segundo color:



Continuamos pintando los vecinos de F con el color opuesto al color de F:



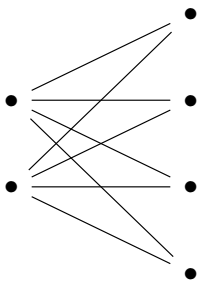
Y ahora pintamos los vecinos de B con el color opuesto a B:



Como hemos logrado pintar todos los vértices del grafo de forma que las aristas siempre conecten vértices de color distinto, concluimos que el grafo es bipartito.

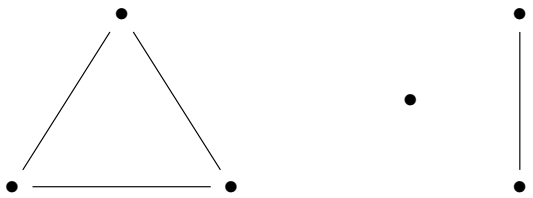
Definición El grafo bipartito completo de m y n, denotado $K_{m,n}$ vértices es el grafo simple donde el conjunto de vértices puede particionarse en V_1 con m vértices y V_2 con n vértices, y donde el conjunto de aristas consiste en todas las aristas de la forma (v_1, v_2) con $v_1 \in V_1$ y $v_2 \in V_2$.

Se muestra como ejemplo el grafo $K_{2,4}$.



6. Más conceptos sobre grafos

Definición Un **grafo conexo** es un grafo donde, dado cualesquiera dos vértices, siempre existe un camino que comienza en uno y termina en el otro. Todos los ejemplos que grafos que hemos visto hasta ahora son conexos. A continuacion, se da un ejemplo de un grafo **no conexo**.



Intuitivamente, un grafo conexo es aquel "de una sola pieza". Mientras que los grafos no conexos parecen, a simple vista, estar formado por varias partes. A cada una de estas "partes"se la llama **componente conexa**. Se nota con $\kappa(G)$ a la cantidad de componentes conexas del grafo. En el ejemplo, $\kappa(G) = 3$.

Definición Sea $G = (V, E)$ un grafo. $G' = (V', E')$ es un **subgrafo** de G si se cumple:

1. $V' \subseteq V$
2. $E' \subseteq E$
3. Para toda arista $e \in E'$, si e incide en los vértices v y w, entonces $v, w \in V'$

Esta definición nos dice que un grafo es subgrafo de otro cuando esta compuesto por algunos de sus vértices (posiblemente todos), algunas de sus aristas (posiblemente todas) y añadiendo la condición especial de que todas las aristas del subgrafo deben tener sus correspondientes vértices en el subgrafo.

Definición Sea $G = (V, E)$. Si $U \subseteq V$, el **subgrafo de G inducido por** U es el subgrafo cuyo conjunto de vértices es U y que contiene todas las aristas de G de la forma (v, w) donde $v, w \in U$.

Definición Un subgrafo G' de un grafo $G = (V, E)$ es un **subgrafo inducido** si existe un conjunto $U \subset V$ tal que G' es el subgrafo de G inducido por U.

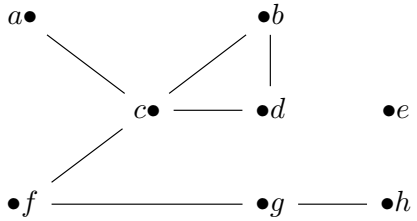
Definición Sea v un vértice en un grafo (dirigido o no dirigido) $G = (V, E)$. El subgrafo de G denotado por $G - v$ tiene el conjunto de vértices $V' = V - \{v\}$ y el conjunto de aristas $E' \subset E$, donde E' contiene todas las aristas en E, excepto aquellas que son incidentes en el vértice v.

Observación El grafo $G - v$ es el subgrafo inducido por $V - \{v\}$.

Definición De manera similar, si e es un arista en un grafo (dirigido o no dirigido) $G = (V, E)$, el subgrafo $G - e$ es el subgrafo que contiene todas las aristas de G, excepto e, y el mismo conjunto de vértices.

Definición Sea G un grafo dirigido o no dirigido. Para cada vertice v de G, el **grado** de v, denotado por $\delta(v)$ es el número de aristas incidentes en v.

Por ejemplo, en el siguiente grafo, $\delta(a) = 1$, $\delta(b) = 2$, $\delta(c) = 4$, $\delta(d) = 2$, $\delta(e) = 0$, $\delta(f) = 2$, $\delta(g) = 2$ y $\delta(h) = 1$



Si $G = (V, E)$ es un grafo no dirigido, luego $\sum_{e \in V} \delta(v) = 2|E|$. Esto se debe a que cada arista de la forma $(a, b) \in E$ suma 1 al grado de a y uno al grado de b, es decir, suma 2 a la sumatoria $\sum_{e \in V} \delta(v)$. Entonces $2|E|$ es el resultado de sumar los grados de todos los vértices en V. Esto implica que el número de vértices de grado impar debe ser par.

7. Algoritmo de Dijkstra

Un problema muy recurrente es encontrar el camino más corto entre dos vértices dados (es decir, un camino que tiene la longitud mínima entre todas los posibles caminos entre esos dos vértices.)

Edsger W. Dijkstra (1930-2002) fue un científico de la computación holandés que ideó un algoritmo que resuelve el problema de la ruta más corta de forma óptima. Además, fue de los primeros en proponer la programación como una ciencia. Ganador del premio Turing en 1972. El premio Turing es "el Nobel de la programación". Poco después de su muerte recibió también el premio de la ACM[2]. El premio pasó a llamarse Premio Dijkstra el siguiente año en su honor.

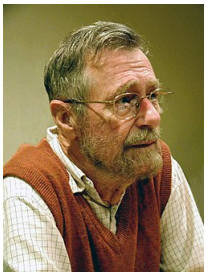


Figura 9: Edsger Wybe Dijkstra

Este algoritmo encuentra la longitud de una ruta más corta del vértice a al vértice z en un grafo $G = (V, E)$ ponderado y conexo. El peso de la arista (i, j) es $w(i, j) > 0$ (es decir, requerimos que los pesos sean siempre positivos). Es importante verificar que estas condiciones se cumplen antes de aplicar el algoritmo.

Como el objetivo es encontrar la ruta de longitud mínima, el algoritmo Dijkstra funciona asignando **distancias** a los vértices, que siempre son relativas al vértice inicial. Notaremos a la distancia del vértice a al vértice v como $D(v)$ (no es necesario mencionar a a porque es fijo durante una ejecución del algoritmo). Algunas distancias son temporales y otras son permanentes. Al ilustrar el algoritmo, por lo general se encierran en un círculo, o se pintan de un color distinto los vértices cuya distancia calculada ya es permanente. Llamaremos T al conjunto de vértices que aún no fueron visitados (y, por lo tanto, tienen una distancia temporal). Esos son los vértices con lo que "tenemos que seguir trabajando". La idea es que, una vez $D(v)$ sea la etiqueta definitiva de un vértice v, $D(v)$ sea la distancia más corta desde a hasta v.
Al inicio, todos los vértices estarán en la lista de no visitados. Cada iteración del algoritmo visita un nuevo nodo, determinando su distancia definitiva. El algoritmo termina cuando z es visitado. Cuando llega ese momento, $D(z)$ es el valor de la distancia del camino más corto desde a hasta z. Los pasos detallados son los siguientes:

1. Primero, notemos que la ruta más corta del vértice a al vértice a, es la ruta de cero aristas, que tiene peso 0. Así, inicializamos $D(a) = 0$.

2. No sabemos aún el valor de una ruta más corta de a a los otros vértices, entonces para cada vértice $v \neq a$, inicializamos $D(v) = \infty$.

3. Inicializamos el conjunto T como el conjunto de todos los vértices, i.e. $T = V$.

4. Seleccionamos un vértice $v \in T$ tal que $D(v)$ sea mínimo.

5. Quitamos el vértice v del conjunto T: $T = T - v$

6. Para cada $t \in T$ adyacente a v, actualizamos su etiqueta: $D(t) = \text{MIN}\{D(t), D(v) + w(v, t)\}$

7. Si $z \in T$, repetimos desde el paso 4, si no, hemos terminado y $D(z)$ es el valor de la ruta mas corta entre a y z.

Ejemplo. En la Figura 10 se muestra un ejemplo de la ejecución del algoritmo de Dijkstra en 10 pasos. El objetivo es encontrar la longitud de la ruta más corta desde el vértice a hacia el vértice e (en este caso, 6). En violeta se marcan los nodos que son visitados en cada paso, mientras que en

[2]Association for Computing Machinery

verde se señalan las aristas que se consideran para (quizás) actualizar las distancias hacia los nodos adyacentes.

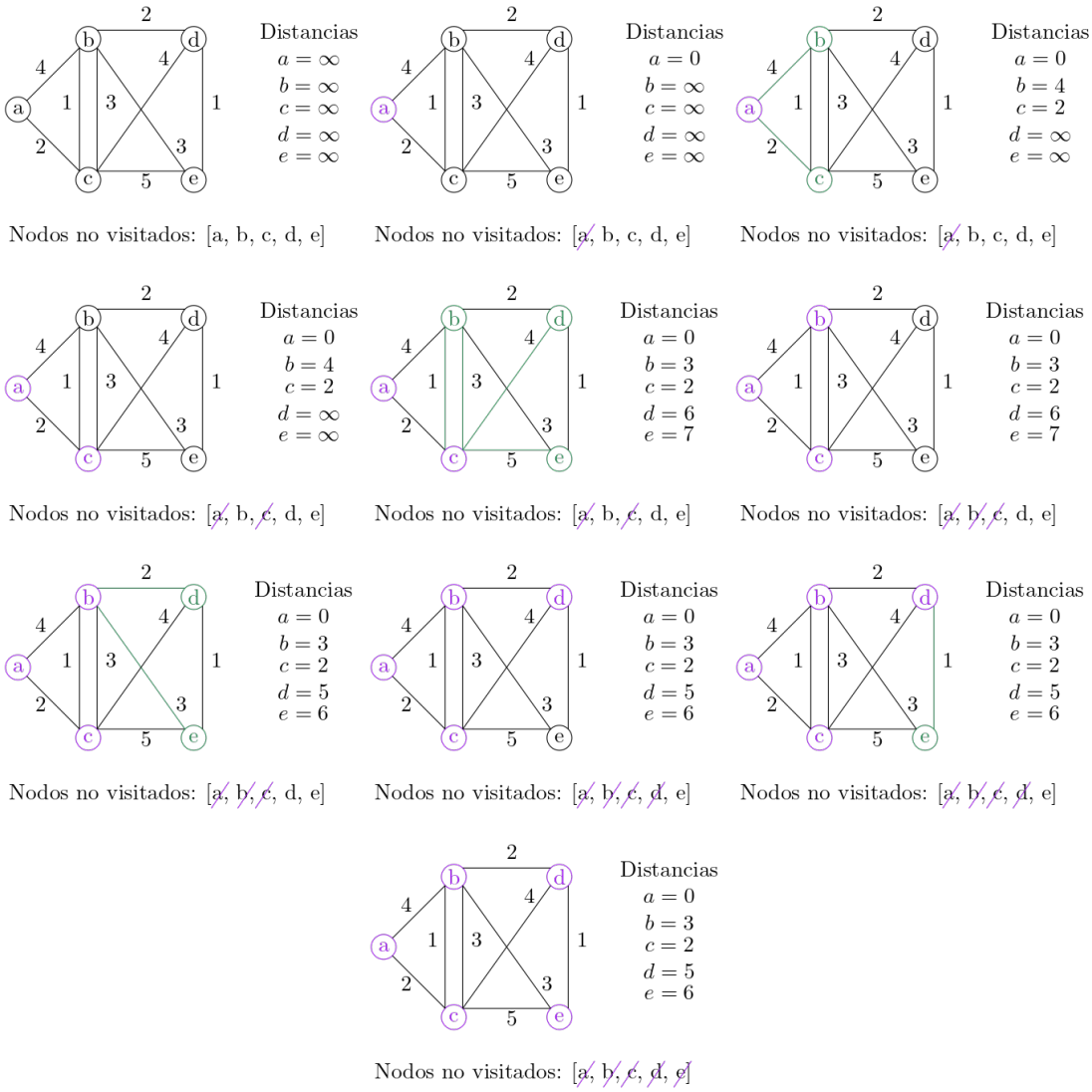


Figura 10: Ejecución del algoritmo de Dijkstra

8. Árboles

Los árboles son un tipo especial de grafo. La definición que dimos en unidades anteriores se puede expresar ahora en términos de teoría de grafos.

Definición. Un **árbol** (libre) T es una grafo simple que satisface lo siguiente: si v y w son vértices en T existe un camino único de v a w.

Recordemos también que llamamos **árbol con raíz** a un árbol donde se ha designado un nodo especial, llamado raíz.

Ahora que sabemos teoría de grafos podemos, además, dar algunas propiedades adicionales de árboles.

Propiedad Un grafo conexo y sin ciclos es un árbol. En efecto, como el grafo es conexo entre dos vértices cualesquiera siempre hay un camino.

Propiedad Un grafo conexo con n vértices y $n-1$ aristas es necesariamente un árbol.

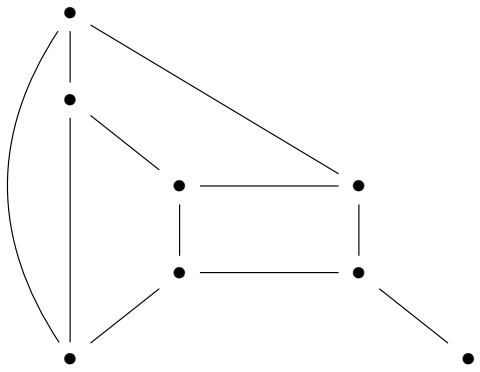
Propiedad Si un grafo con n vértices y $n-1$ aristas, no contiene ciclos entonces necesariamente es un árbol.

8.1. Árboles de expansión

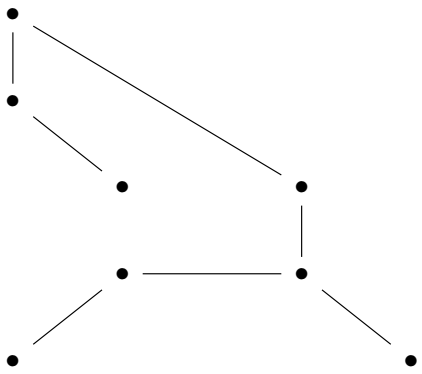
Se dice que árbol T es un **árbol de expansión** de un grafo G si:

1. T es un subrafo de G.
2. T es un árbol.
3. T contiene todos los vértices de G.

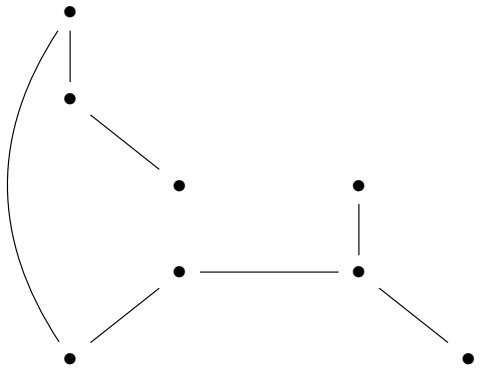
Por ejemplo, dado este grafo



Este es un árbol de expansión de ese grafo:



Nota El árbol de expansión de un grafo en general no es único, por ejemplo, acá mostramos otro árbol de expansión distinto del mismo grafo.



- Un árbol de expansión para un grafo existe si y solo si el grafo es conexo.

En efecto, todos los árboles pueden pensarse como grafos conexos sin ciclos. Si un grafo G tiene un árbol de expansión T, entonces entre dos vértices cualesquiera existe un camino en el árbol que los une (pues todos los árboles son conexos). Ahora bien, como T es subgrafo de G, necesariamente debe existir ese mismo camino en el grafo G.
El algoritmo de búsqueda en profundidad, normalmente llamado DFS por sus siglas en inglés *Depth First Search* permite encontrar el árbol de expansión de un grafo conexo. Este algoritmo puede verse como versión generalizada del recorrido en pre-orden. La estrategia consiste en partir de un vértice determinado v y a partir de allí, cuando se visita un nuevo vértice, explorar cada camino que salga de él. Un camino deja de explorarse cuando se llega a un vértice ya visitado. Si existen vértices no alcanzables, el recorrido queda incompleto; entonces, se debe "volver hacia atrás"hasta encontrar un vértice donde hayamos dejado un camino sin explorar, y repetir el proceso.

Se utiliza, además, un orden sobre los vértices para desempatar en casos donde tengamos más de una opción. Normalmente, como convención interna, utilizaremos el orden alfabético cuando los nodos lleven letras por nombres, o el orden natural cuando los vértices sean numerados. Dado un orden $v_1, v_2, ..., v_n$ de los vértices, el algoritmo puede expresarse formalmente como sigue:

1. Dado el grafo $G = (V, E)$, inicializamos el árbol $T = (V', E')$ con $V' = \{v_1\}$ y $E' = \emptyset$. La idea sera ir agregando aristas a E' a medida que lo necesitemos. Definimos además una variable $w = v_1$ que llevará el nodo donde estamos parados actualmente.

2. Mientras exista v tal que (w, v) es una arista que al agregarla a T no genera un ciclo, realizamos lo siguiente:

   a) Elegimos la arista (w, v_k) con k mínimo tal que al agregarla a T no genera un ciclo.

   b) Agregamos la arista (w, v_k) a E'.

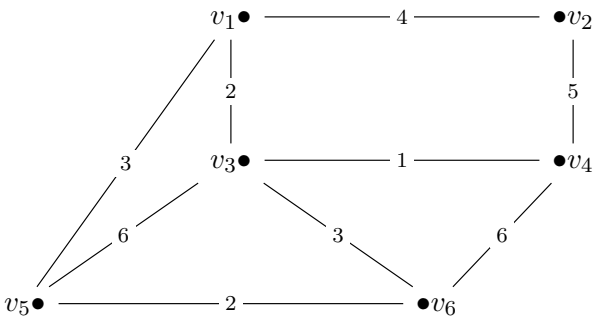
   c) Agregamos v_k a V'

   d) Actualizamos $w = v_k$

3. Si $V' = V$ hemos terminado y T es un árbol de expansión del grafo G. Si $w = v_1$, el grafo es disconexo, y por lo tanto jamás podremos encontrar un árbol de expansión para el mismo. Si no se da ninguna de las dos situaciones, actualizamos el valor de w para que sea el padre de w en el árbol T, y repetimos desde el paso 2. Dar este paso hacia atrás nos obligará a explorar otros caminos.

8.2. Árbol de expansión mínima

Ejemplo El grafo con pesos de la figura muestra seis ciudades y los costos de construir carreteras entre ellas. Se desea construir el sistema de carreteras de menor costo que conecte a las seis ciudades. La solución debe necesariamente ser un árbol de expansión ya que debe contener a todos los vértices y para ser de costo mínimo, sería redundante tener dos caminos entre ciudades. Entonces lo que necesitamos es el árbol de expansión del grafo que sea de peso mínimo.



Definición Sea G un grafo con pesos. Un árbol de expansión mínima de G es un árbol de expansión de G que tiene peso mínimo entre todos los posibles.

Nota. El algoritmo DFS no asegura que el árbol encontrado sea de peso mínimo.

El algoritmo de Prim permite encontrar un árbol de expansión mínimo para un grafo con pesos conexo de vértices $v_1, v_2, ...v_n$.

El algoritmo incrementa continuamente el tamaño de un árbol, de manera similar al algoritmo DFS, comenzando por un vértice inicial al que se le van agregando sucesivamente vértices cuya distancia a los anteriores es mínima. Esto significa que en cada paso, las aristas a considerar son aquellas que inciden en vértices que ya pertenecen al árbol. El árbol está completamente construido cuando no quedan más vértices por agregar. Se puede describir formalmente como sigue:

1. Inicializamos un conjunto de vértices agregados V' y un conjunto de aristas agregadas E' ambos como el conjunto vacío.

2. Elegimos un vértice inicial arbitrario, y lo agregamos a V'.

3. Si V' contiene todos los vértices del árbol, entonces E' conforma el árbol de expansión mínima y terminamos el algoritmo. Si no, continuamos con el siguiente paso.

4. Elegimos, entre todas las aristas (v_i, v_j) con $v_i \in V'$ y $v_j \notin V'$, la que tiene peso mínimo[3]. Añadimos dicha arista a E' y agregamos al vértice v_j al conjunto V'.

5. Repetimos el paso 3 al 5 hasta que el algoritmo finaliza.

Ejemplo. En la Figura 11 se muestra un ejemplo de la ejecución del algoritmo Prim en 12 pasos. El objetivo es encontrar un árbol de expansión mínimo. En violeta se marcan los nodos y aristas que son agregados al árbol en cada paso, mientras que en naranja se señalan las aristas candidatas que van surgiendo.

[3]Notemos que el grafo es no dirigido, con lo cual, la arista (v_i, v_j) es equivalente a la arista (v_j, v_i)

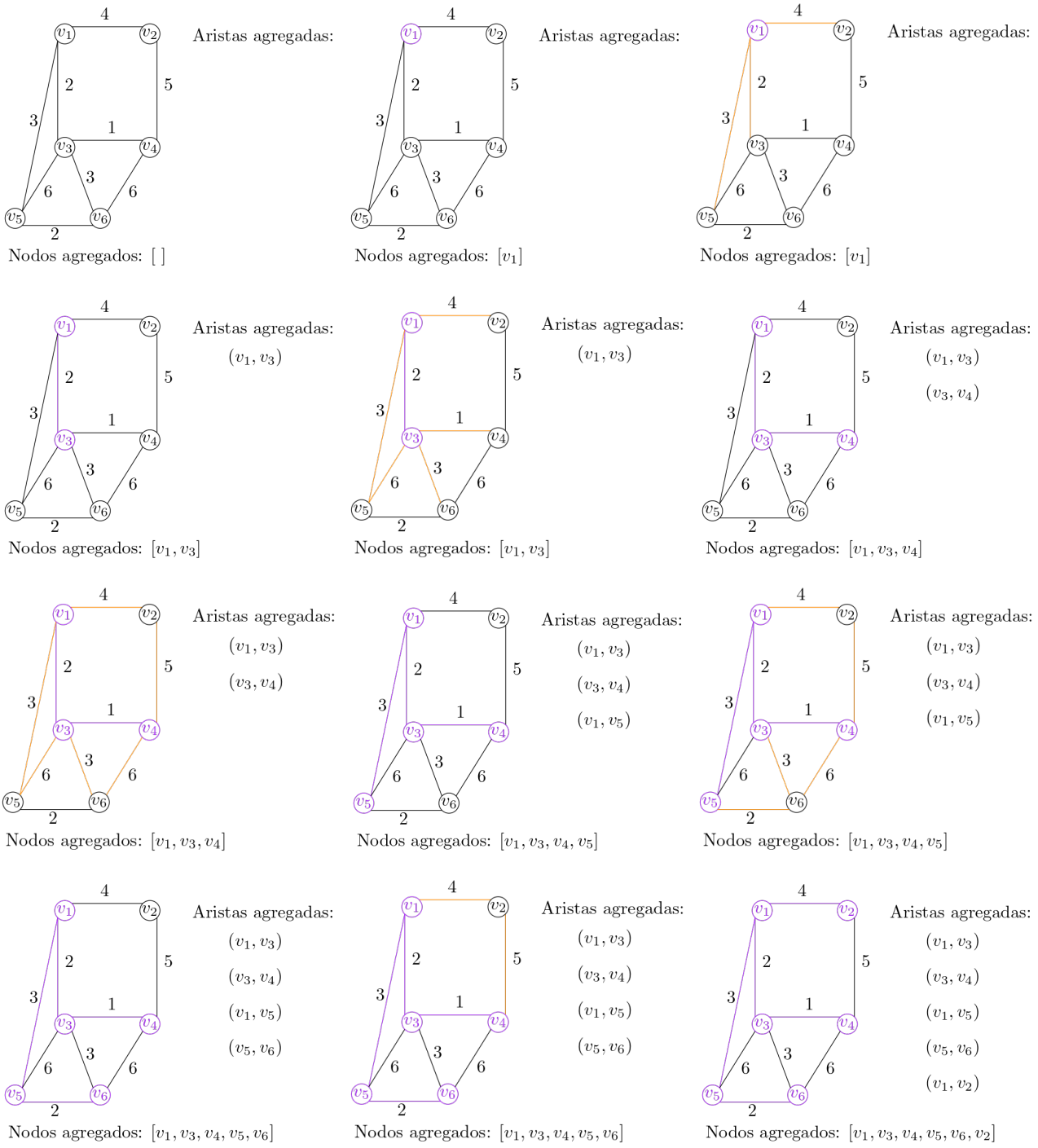


Figura 11: Ejecución del algoritmo Prim

9. Introducción a `networkx`

NetworkX es un paquete de Python para crear, manipular y estudiar la estructura de grafos complejos, que trae incorporada distintas funcionalidades para grafos (incluyendo los algoritmos estudiados). A continuación se presentan ejemplos de uso.

```
import networkx as nx

# Creamos un grafo vacío
G = nx.Graph()

# Imprimimos el grafo
print(G) # Graph with 0 nodes and 0 edges

# Agregamos un nodo individual
G.add_node(1)

# Agregamos nodos desde una lista
G.add_nodes_from([2, 3])

print(G) # Graph with 3 nodes and 0 edges

# Agregamos aristas individuales
G.add_edge(1, 2)
e = (2, 3)
G.add_edge(*e)

# Agregamos aristas desde una lista
G.add_edges_from([(1,1), (1, 3)])

# Podemos imprimir la cantidad de nodos
G.number_of_nodes() # 3
# Podemos imprimir la cantidad de aristas
G.number_of_edges() # 4

# Si agregamos aristas de nodos no existentes, los nodos son agregados
G.add_edge(4, 5)
G.number_of_nodes() # 5
G.number_of_edges() # 5

# Podemos imprimir las listas de nodos y aristas
list(G.nodes) # [1, 2, 3, 4, 5]
list(G.edges) # [(1, 2), (1, 1), (1, 3), (2, 3), (4, 5)]

# Podemos obtener un diccionario con el grado de cada vértice
dict(G.degree) # {1: 4, 2: 2, 3: 2, 4: 1, 5: 1}

# También es posible eliminar nodos o aristas
G.remove_node(4) # Eliminará también las aristas incidentes
G.remove_nodes_from([3,5]) # Eliminará también las aristas incidentes
G.remove_edge(1,1)

list(G.nodes) # [1, 2]
list(G.edges) # [(1, 2)]
```

También es posible crear grafos con pesos de manera muy sencilla. Para lograrlo, simplemente se debe ingresar el valor del peso de la arista como un tercer argumento. Al crear grafos ponderados, es posible ejecutar sobre ellos algoritmos como Dijkstra o Prim. Se presenta un ejemplo a continuación.

```
import networkx as nx

# Creamos un multigrafo
G = nx.Graph()
G.add_nodes_from(['a', 'b', 'c', 'd', 'e'])
G.add_weighted_edges_from([('a', 'b', 4),
                           ('a', 'c', 2),
                           ('b', 'c', 1),
                           ('b', 'd', 2),
                           ('b', 'e', 3),
                           ('c', 'd', 4),
                           ('c', 'e', 5),
                           ('d', 'e', 1)])
# Calculamos el camino de longitud mínima desde 'a' hasta 'e'
nx.dijkstra_path(G, source='a', target='e') # ['a', 'c', 'b', 'e']
# Calculamos la longitud de dicho camino
nx.dijkstra_path_length(G, source='a', target='e') # 6

# Creamos un grafo
G = nx.Graph()
G.add_nodes_from(['v1', 'v2', 'v3', 'v4', 'v5', 'v6'])
G.add_weighted_edges_from([('v1', 'v2', 4),
                           ('v1', 'v3', 2),
                           ('v1', 'v5', 3),
                           ('v2', 'v4', 5),
                           ('v3', 'v4', 1),
                           ('v3', 'v5', 6),
                           ('v3', 'v6', 3),
                           ('v4', 'v6', 6),
                           ('v5', 'v6', 2)])
# Calculamos el árbol de expansión mínimo
mst = nx.minimum_spanning_tree(G, algorithm='prim')

# Calculamos el peso del árbol
mst.size(weight='weight') # 12.0
# Imprimimos las aristas
list(mst.edges) # [('v1', 'v3'), ('v1', 'v2'), ('v3', 'v4'), ('v3', 'v6'),
# ('v5', 'v6')]
```

Referencias

[] Tutorial oficial de NetworkX https://networkx.org/documentation/stable/tutorial.html
Lectura recomendada

[1] Johnsonbaugh *Matemáticas discretas*. 6ta Edición. Capítulos 8 y 9

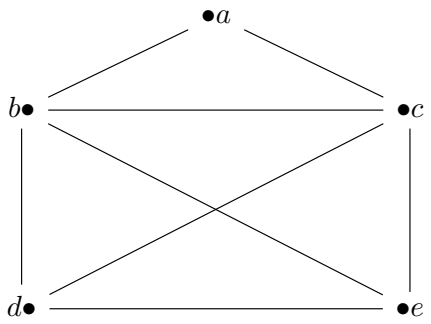
[2] Grimaldi, R. *Matemáticas Discretas y Combinatoria.*

Advertencia. La terminología asociada a la teoría de grafos no se ha estandarizado aún. Al leer artículos y libros sobre grafos, es necesario verificar las definiciones que se emplean. Ante cualquier duda, consultar con los docentes de la cátedra.

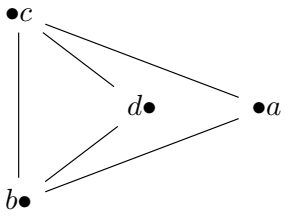
10. Ejercicios

1. Para cada uno de los siguientes grafos, decidir si existe un camino del vértice a al vértice a que pase por cada arista exactamente una vez.

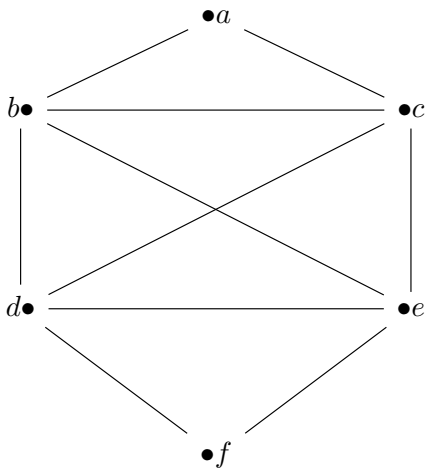
a)



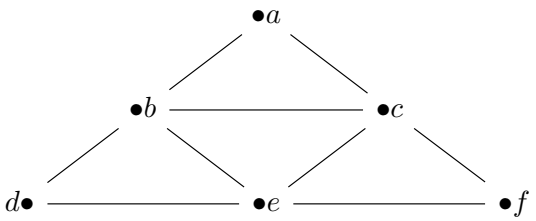
b)



c)



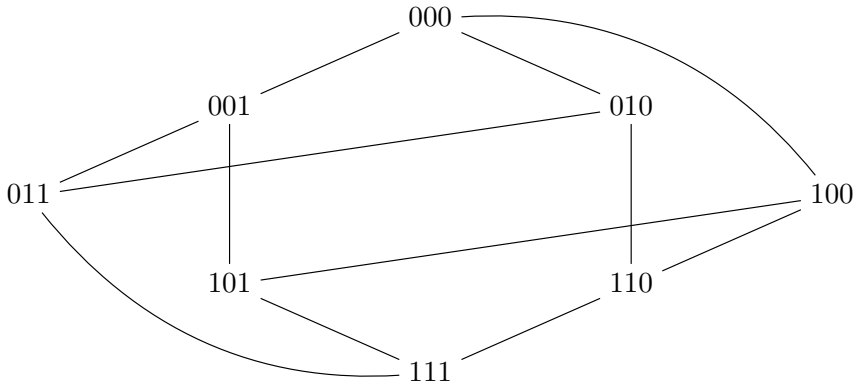
d)



2. Dibujar los grafos K_6 y K_7.

3. Encontrar una fórmula para la cantidad de aristas del grafo K_n (en función de n).

4. Dado un número d, consideramos un conjunto de nodos compuesto por todas las secuencias binarias de exactamente d dígitos y dibujamos una arista entre ellos si y solo si ambas secuencias difieren en exactamente un dígito. A este grafo lo llamamos **hipercubo de d dimensiones**. Por ejemplo, para $d = 3$, obtenemos:

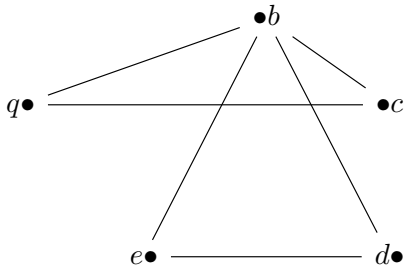


.

   a) Determine el grado de cada vértice, como una función de d.

   b) Determine la cantidad de vértices y aristas en el grafo, como una función de d.

5. ¿Es posible dibujar el siguiente grafo de forma que sus aristas no se crucen entre si?



6. En cierto edificio de la ciudad, los vecinos se llevan bastante mal entre ellos. Sabemos que el edificio es de 25 departamentos, y en cada departamento vive una familia. Cada familia se lleva bien con exactamente otras 5 familias del edificio. ¿Es posible esta situación?

7. Definimos el doble de un grafo como dos copias de dicho grafo con aristas adicionales uniendo los vértices correspondientes. La mejor forma de verlo es con un ejemplo. En la siguiente figura se puede ver un grafo A y su doble, el grafo B.

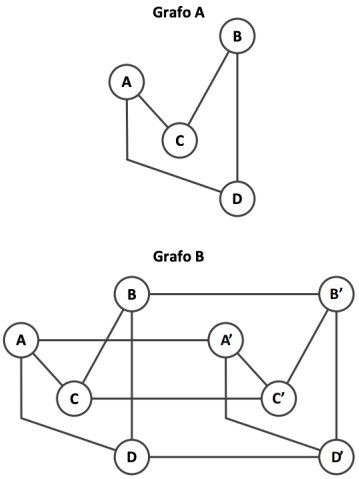
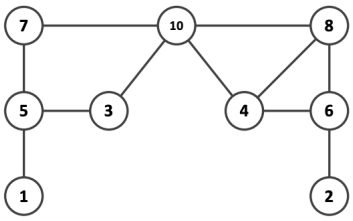


Figura 12: Grafos dobles

a) ¿Es el grafo A bipartito?

b) ¿Es el grafo B bipartito?

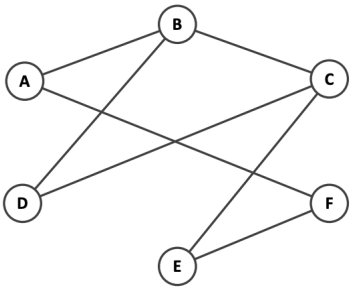
8. En la siguiente figura representamos un serie de edificios en Venecia, junto con puentes que los unen:



a) ¿Es el grafo bipartito?

b) Un inspector de puentes, que vive en el edificio número 10, esta encargado de revisar el estado de los puentes que se ven en la figura. ¿Existe forma de que recorra cada puente exactamente una vez, partiendo de su edificio y regresando al mismo? Justifique.

9. Dado G = (V, E), diremos que la **distancia** de un vértice a otro es la longitud del menor camino posible entre ambos vértices. El **diámetro** de un grafo es la mayor distancia entre todos los pares de vértices del mismo. Dado el siguiente grafo:



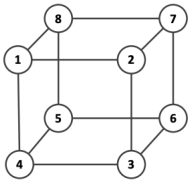
a) ¿Cual es la distancia de B a F?

b) Encuentre el diámetro del grafo.

c) Decida si existe un camino que pase por cada vértice exactamente una vez.

d) Decida si existe un camino que pase por cada arista exactamente una vez.

10. Observe el siguiente grafo:



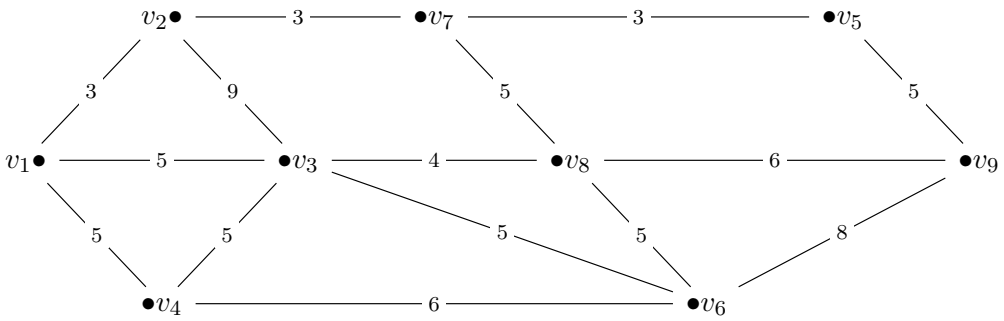
a) Decida si existe un camino que pase por cada vértice exactamente una vez.

b) Decida si existe un camino que pase por cada arista exactamente una vez.

c) ¿Es el grafo bipartito?

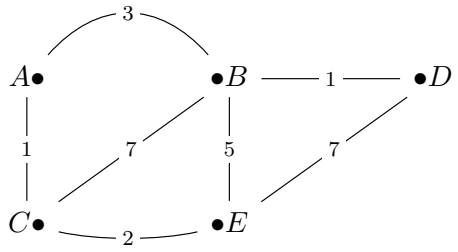
d) ¿Cual es el diametro de este grafo?

11. Para el siguiente grafo, encontrar la longitud de la ruta más corta del vértice v_1 al vértice v_9 utilizando el algoritmo de Dijkstra y el árbol de expansión mínimo utilizando el algoritmo de Prim. Describir de manera detallada los algoritmos y sus pasos.

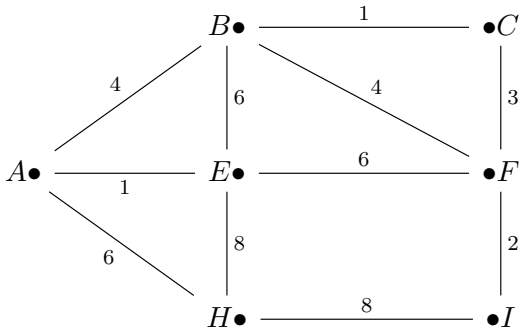


12. Para el siguiente grafo, encuentre la longitud de la ruta más corta del vértice C al vértice E utilizando el algoritmo de Dijkstra y el árbol de expansión mínimo utilizando el algoritmo de

Prim. Describir de manera detallada los algoritmos y sus pasos.



13. Para el siguiente grafo, encontrar la longitud del camino más corto del vértice A al vértice I, utilizando el algoritmo de Dijkstra. Describir de manera detallada el algoritmo y sus pasos.



14. Describa que es un árbol de expansión de un grafo. ¿Qué propiedad debe cumplirse para que el árbol de expansión exista? ¿Cuando podemos decir que el árbol de expansión del grafo es mínimo?