



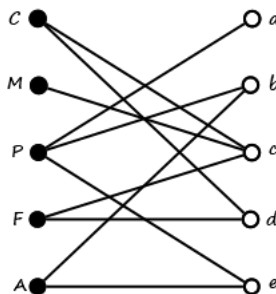
## Teórica 11: Matching, conjunto independiente y recubrimientos

### 1. Introducción

Analicemos algunas situaciones:

**Ejemplo 1.** Hay un conjunto de cinco personas y un conjunto de 5 trabajos para realizar. Sean las personas Carlos, Marcela, Pedro, Fernando y Andrea, y los trabajos a, b, c, d y e. Carlos está capacitado para realizar los trabajos c y d, Marcela para c, Pedro para a, b y e, Fernando para c y d, y Andrea para b y e. ¿Es posible realizar una distribución del trabajo de modo que se puedan realizar todos los trabajos simultáneamente?

Podemos modelar el problema mediante un grafo  $G = (V, X)$  donde el conjunto de vértices representa a las personas y los trabajos, y hacemos un vértice representando a una persona adyacente a los vértices correspondientes a los trabajos para los que está capacitado. Obtenemos el siguiente grafo:

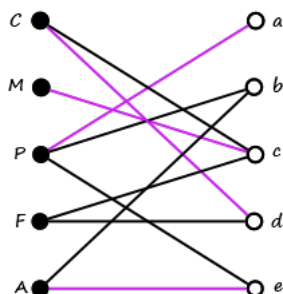


Para poder realizar todos los trabajos simultáneamente, debemos asignarle exactamente un trabajo a cada persona. Intentemos hacerlo:

- El trabajo a sólo puede realizarlo Pedro, así que no tenemos elección: le asignamos a Pedro el trabajo a.
- Marcela sólo sabe realizar el trabajo c, tampoco tenemos elección: le asignamos a Marcela el trabajo c.
- Carlos está capacitado para realizar los trabajos c y d, pero el c ya lo tiene asignado Marcela: le asignamos el trabajo d a Carlos.
- Fernando puede realizar los trabajos c y d, pero ambos trabajos ya están asignados, por lo que no tenemos trabajo para asignarle Fernando.

Esto ya muestra que no será posible realizar todos los trabajos simultáneamente.

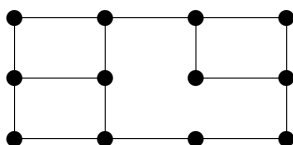
Andrea puede realizar los trabajos b o e. Ambos trabajos todavía están libres y le podemos asignar alguno de los dos, por ejemplo el trabajo e. Entonces sí podemos realizar simultáneamente 4 de los 5 trabajos.



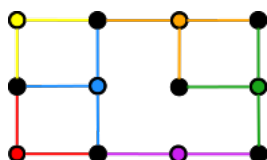
En esta aplicación, que se conoce como problema de asignación de personal, estamos interesados en encontrar un conjunto de aristas de cardinal máximo, tal que no haya dos aristas en el conjunto que incidan sobre el mismo vértice.

Analicemos otro escenario.

**Ejemplo 2.** El grafo de la figura representa el mapa de una ciudad. Se quiere ubicar policías en las esquinas de modo que todas las cuadras estén bajo vigilancia, o sea, cada cuadra tiene que tener un policía al menos en una de las esquinas. ¿Cuál es el mínimo número de policías necesarios?



Si ubicamos un policía en cada uno de los vértices coloreados, todas las cuadras estarán cubiertas.

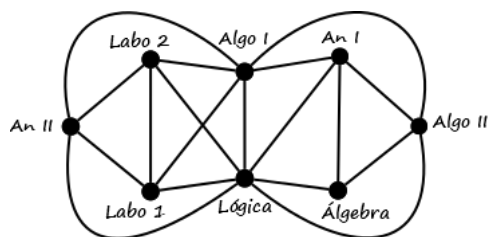


En este caso estamos necesitando 6 policías. Pero, ¿habrá forma de hacerlo con 5?

Ahora nuestro objetivo es encontrar un conjunto de vértices de cardinal mínimo, tal que toda arista del grafo incida, al menos, sobre un vértice de este conjunto.

Volvamos al problema de las aulas de la clase de coloreo, pero con un objetivo distinto.

**Ejemplo 3.** Supongamos que no tenemos restricción en cantidad de aulas y que el DC equipó una sala (sólo una por ahora) con la última tecnología y quiere que sea utilizada por la mayor cantidad de materias posible. Volvamos a dibujar el grafo que representa nuestra instancia:





Si le asignamos la sala a la materia Lógica, ninguna otra materia podrá usarla, porque Lógica se solapa en horario con todas las demás.

En cambio, si le asignamos la sala a Álgebra, también la podría utilizar por ejemplo Algo I (o Labo 1, o Labo 2, o Análisis II), pero no Algo II ni Análisis I ni Lógica. En este caso, la sala sería utilizada por dos materias. Pero, ¿será posible que la sala sea utilizada por tres o más materias?

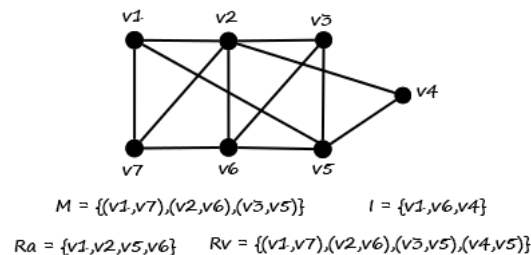
Lo que queremos hallar en este ejemplo es un conjunto de vértices de cardinal máximo tal que no haya dos vértices adyacentes dentro del conjunto.

Este tipo de escenarios dan origen a las siguientes definiciones y estudios.

**Definición 1.** Dado un grafo  $G = (V, X)$

- Un **matching o correspondencia** entre los vértices de  $G$ , es un conjunto  $M \subseteq X$  de aristas de  $G$  tal que para todo  $v \in V$ ,  $v$  es incidente a lo sumo a una arista  $e \in M$ .
- Un **conjunto independiente** de vértices de  $G$ , es un conjunto de vértices  $I \subseteq V$  tal que para toda arista  $e \in X$ ,  $e$  es incidente a lo sumo a un vértice  $v \in I$ .
- Un **recubrimiento de las aristas** de  $G$ , es un conjunto  $R_a \subseteq X$  de aristas tal que para todo  $e \in X$ ,  $e$  es incidente al menos a un vértice  $v \in R_a$ .
- Un **recubrimiento de los vértices** de  $G$ , es un conjunto  $R_v \subseteq X$  de aristas tal que para todo  $v \in V$ ,  $v$  es incidente al menos a una arista  $e \in R_v$ .

**Ejemplo 4.**



Los problemas de optimización relacionados a estos conceptos son:

**Definición 2.** Dado un grafo  $G = (V, X)$

- El problema de **matching máximo** consiste en encontrar un matching de cardinal máximo entre todos los matchings de  $G$ .
- El problema de **conjunto independiente máximo** consiste en encontrar un conjunto independiente de cardinal máximo entre todos los conjuntos independientes de  $G$ .
- El problema de **recubrimiento de aristas mínimo** consiste en encontrar un recubrimiento de aristas de cardinal mínimo entre todos los recubrimientos de aristas de  $G$ .
- El problema de **recubrimiento de vértices mínimo** consiste en encontrar un recubrimiento de vértices de cardinal mínimo entre todos los recubrimientos de vértices de  $G$ .



Nuevamente, aunque parezcan problemas muy parecidos, unos sobre vértices y otros sobre aristas, computacionalmente son muy distintos. Para el problema de matching máximo existen algoritmos polinomiales para resolverlo, mientras que para conjunto independiente máximo no se conoce ninguno (y se piensa que no existe).

Podemos enunciar una relación interesante entre dos de estos conceptos:

**Lema 1.** Sea  $G = (V, X)$  y  $S \subseteq V$ .  $S$  es un conjunto independiente  $\iff V \setminus S$  es un recubrimiento de aristas.

*Demostración.*  $\implies$ ) Por el absurdo, supongamos que  $S$  es un conjunto independiente y  $V \setminus S$  no es un recubrimiento de aristas. Entonces existe una arista  $(u, v) \in X$  no recubierta, es decir,  $u \notin V \setminus S$  y  $v \notin V \setminus S$ . Ésto implica que  $u \in S$  y  $v \in S$ . Esto contradice que  $S$  es conjunto independiente, ya que  $u$  y  $v$  son adyacentes.

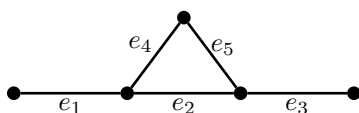
$\impliedby$ ) Ahora, también por el absurdo, supongamos que  $V \setminus S$  es un recubrimiento de aristas y  $S$  no es un conjunto independiente. Entonces existen dos vértices  $u$  y  $v$  en  $S$  adyacentes. Ésto implica que  $u \notin V \setminus S$  y  $v \notin V \setminus S$ . Pero entonces la arista  $(u, v)$  no está cubierta por  $V \setminus S$ , contradiciendo que es un recubrimiento de aristas.

■

**Corolario 1.** Sea  $G = (V, X)$  y  $S \subseteq V$ .  $S \subseteq V$  es un conjunto independiente máximo  $\iff V \setminus S$  es un recubrimiento de aristas mínimo.

Este Lema implica que para el problema de aristas mínimo tampoco se conocen algoritmos polinomiales.

Por el contrario, esta relación no se mantiene entre matchings y recubrimientos de vértices, como lo muestra este ejemplo:



$$M = \{e_3, e_4\}$$

$$X \setminus M = \{e_1, e_2, e_5\}$$

no es recubrimiento

$$R = \{e_1, e_3, e_4\}$$

$$X \setminus R = \{e_2, e_5\}$$

no es matching

Sin embargo, veremos que existe una relación entre ambos problemas.

## 2. Relación entre coloreo y conjunto independiente

Ahora que conocemos la definición de conjunto independiente podemos definir el problema de coloreo de vértices desde un enfoque equivalente:

- Cada conjunto de vértices pintados con el mismo color es un conjunto independiente.
- Entonces, podemos definir un coloreo de  $G$  como una partición de sus vértices en conjuntos independientes.
- El número cromático de  $G$ , es la cantidad de conjuntos de la partición con menos conjuntos.

No es cierto que en un coloreo óptimo de  $G$  los vértices de un conjunto independiente de cardinal máximo estén siempre pintados todos del mismo color. Sin embargo, una heurística golosa para colorear un grafo es (recordar que  $G[W]$  indica el subgrafo inducido de  $G$  por los vértices de  $W$ ):



```
coloreoPorCI( $G$ )  
  entrada:  $G = (V, X)$   
  salida:  $f$  coloreo de los vertices de  $G$   
  
   $i \leftarrow 0$   
  mientras  $V \neq \emptyset$   
     $i \leftarrow i + 1$   
     $S \leftarrow \text{conjIndepMaximal}(G_{[V]})$   
    para todo  $v \in S$  hacer  
       $f[v] \leftarrow i$   
    fin para  
     $V \leftarrow V \setminus S$   
  retornar  $f$ 
```

Notar que en cada iteración del ciclo,  $S$  es un conjunto independiente maximal, no necesariamente máximo. Encontrar un conjunto independiente maximal es fácil computacionalmente, mientras que encontrar uno máximo no lo es.

Un procedimiento constructivo goloso para encontrar un conjunto independiente maximal *grande* de un grafo  $G$  podría incorporar en cada iteración el vértice candidato adyacente a la menor cantidad de vértices candidatos, donde un vértice es candidato si no es adyacente a ningún vértice ya incorporado al conjunto. Éste es su pseudocódigo (recordar que  $N(v)$  indica el conjunto de vértices adyacentes a  $v$ ):

```
conjIndepMaximal( $G$ )  
  entrada:  $G = (V, X)$   
  salida:  $I$  conjunto independiente maximal  
  
   $I \leftarrow \emptyset$   
  mientras  $V \neq \emptyset$  hacer  
     $v \leftarrow \arg \min\{|N(v) \cap V| : v \in V\}$   
     $I \leftarrow I \cup \{v\}$   
     $V \leftarrow V \setminus \{u : u \in V \text{ y } (u, v) \in X\}$   
  retornar  $I$ 
```

### 3. Matching máximo

Estudiemos propiedades de este problema que nos conducirán a un algoritmo polinomial para resolverlo.

#### Definición 3.

- Un vértice  $v$  se dice *saturado* por un matching  $M$  si hay una arista de  $M$  incidente a  $v$ .
- Dado un matching  $M$  en  $G$ , un *camino alternado* en  $G$  con respecto a  $M$ , es un camino simple donde se alternan aristas que están en  $M$  con aristas que no están en  $M$ .
- Dado un matching  $M$  en  $G$ , un *camino de aumento* en  $G$  con respecto a  $M$ , es un camino alternado entre vértices no saturados por  $M$ .



El siguiente teorema identifica cuando un matching es máximo.

**Teorema 1.** Dado  $G = (V, X)$ ,  $M$  es un matching máximo de  $G$  si, y sólo si, no existe un camino de aumento en  $G$  con respecto a  $M$ .

Para organizar su demostración, primero vamos a enunciar un lema.

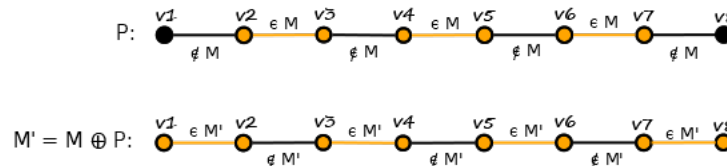
**Lema 2.** Sean  $M_0$  y  $M_1$  dos matching en  $G$  y sea  $G' = (V, X')$  con  $X' = (M_0 - M_1) \cup (M_1 - M_0)$ . Entonces las componentes conexas de  $G'$  son de alguno de los siguientes tipos:

- *vértice aislado*
- *circuito simple con aristas alternadamente en  $M_0$  y  $M_1$*
- *camino simple con aristas alternadamente en  $M_0$  y  $M_1$ .*

*Demostración:* En un matching a lo sumo hay una arista incidente a cada vértice. Como  $X' \subseteq M_0 \cup M_1$ ,  $\forall v \in V$ , en  $X'$  a lo sumo hay una arista de  $M_0$  incidente a  $v$  (porque  $M_0$  es matching) y una de  $M_1$  (porque  $M_1$  es matching), lo que implica que  $d_{G'}(v) \leq 2$  para todo  $v \in V$ . Esto implica que las componentes conexas de  $G'$  son vértices aislados, caminos simples o circuitos simples.

Como en un matching no puede haber dos aristas incidentes sobre el mismo vértice, los caminos y circuitos deben alternar entre aristas de  $M_0$  y de  $M_1$ . ■

*Demostración Teorema:*  $\Rightarrow$ )  $M$  matching máximo. Por el absurdo supongamos que  $\exists P$  camino de aumento en  $G$  con respecto a  $M$ ,  $P : v_0 v_1 v_2 \dots v_{2t+1}$ .  $P$  tiene una arista más en  $X \setminus M$  que en  $M$ . Definimos  $M' = M \oplus P = (M - P) \cup (P - M)$ .



Al hacer  $M - P$ , los vértices intermedios de  $P$  dejan de estar saturados (las puntas no estaban saturadas por ser camino de aumento). Entonces es posible incorporar al matching  $M'$  aristas alternadas de  $P$  (una sí, una no, una sí, ..., una sí). Entonces  $M'$  es un matching y  $|M'| = |M| + 1$ , contradiciendo que  $M$  es máximo.

$\Leftarrow$ )  $G$  no tiene camino de aumento con respecto a  $M$ . Si  $M$  no es máximo,  $\exists M'$  matching con  $|M'| > |M|$ . Sea  $G' = (V, X')$  con  $X' = (M' - M) \cup (M - M') = (M' \cup M) \setminus (M' \cap M)$ . Por el Lema 2, las componentes conexas de  $G'$  son vértices aislados, circuitos pares o caminos. Como  $|M'| > |M|$ ,  $G'$  tiene que tener por lo menos una componente conexa que sea un camino que comienza y termina en aristas de  $M'$ . Esa componente conexa es un camino de aumento en  $G$  con respecto a  $M$ , contradiciendo la hipótesis. ■

Este teorema deriva en el siguiente algoritmo para encontrar un matching de cardinal máximo:



```
matchingMaximo( $G$ )  
  entrada:  $G = (V, X)$   
  salida:  $M$  matching maximo de  $G$   
  
   $M \leftarrow e$  una arista cualquiera  
  mientras exista camino de aumento en  $G$  con respecto a  $M$  hacer  
     $P \leftarrow \text{caminoAumento}(G, M)$   
     $M \leftarrow M \oplus P$   
  fin mientras  
  retornar  $M$ 
```

**Teorema 2.** El algoritmo *matchingMaximo* aplicado a  $G$  retorna un matching de cardinal máximo de  $G$ .

*Demostración.* Llamemos  $M_i$  al valor de la variable  $M$  en la  $i$ -ésima iteración del ciclo del algoritmo. Demostremos primero que  $M_i$  es matching para todo  $i$ . Para ésto haremos inducción en las iteraciones del ciclo.

**Caso base:** Antes de entrar al ciclo,  $k = 0$ ,  $M_0$  está compuesto sólo por una arista. Es un matching trivial.

**Paso inductivo:** Consideremos  $M_i$  de la iteración  $i$ .

Nuestra hipótesis inductiva es:  $M_{i'}$ , para todo  $i' < i$ , es matching.

Como  $M_{i-1}$  es matching por H.I. y  $P$  es un camino de aumento de  $G$  con respecto a  $M_{i-1}$ , por el Teorema 3,  $M_i = M_{i-1} \oplus P$  es matching y  $|M_i| = |M_{i-1}| + 1$ .

El ciclo termina, porque en cada iteración el cardinal de  $M$  aumenta en 1 y no repite aristas. Como hay  $m$  aristas, a lo sumo se puede iterar  $m$  veces.

Ahora ya sabemos que el ciclo termina y al salir  $M$  es matching. Para completar la demostración nos falta ver que este matching es máximo. Cuando se sale del ciclo se cumple que no existe camino de aumento con respecto a  $M$ . Por el Teorema 3, ésto implica que  $M$  es máximo. ■

Nos queda pendiente tratar el problema de encontrar un camino de aumento de un grafo con respecto a un matching. Por su complejidad, lo dejaremos afuera de la materia, lo único que vamos a decir es que existen algoritmos polinomiales para resolverlo [1].

Como el ciclo a lo sumo se realiza  $m$  veces y cada operación dentro del ciclo es polinomial, resulta un algoritmo polinomial para calcular un matching máximo de un grafo.

## 4. Relaciones

Finalmente, analicemos las relaciones entre estos conceptos.



**Teorema 3.** Dado un grafo  $G = (V, X)$  sin vértices aislados, si  $M$  es un matching máximo de  $G$  y  $R_v$  un recubrimiento mínimo de los vértices de  $G$ , entonces  $|M| + |R_v| = n$ .

*Demostración.*  $M$  matching máximo. Sea  $U$  el conjunto de vértices no saturados por  $M$ . Como  $M$  es máximo,  $U$  tiene que ser un conjunto independiente y  $|U| = n - 2|M|$ . Como no hay vértices aislados, podemos elegir una arista incidente a cada vértice de  $U$ , llamamos  $F$  a ese conjunto de aristas. Entonces  $M \cup F$  es un recubrimiento de los vértices de  $G$  con cardinal  $|M| + |F| = |M| + |U| = |M| + n - 2|M|$ . Implicando  $|R_v| \leq n - |M|$  porque  $|R_v|$  es mínimo, es decir,  $|R_v| + |M| \leq n$ .

Ahora  $R_v$  recubrimiento mínimo. Sea  $H = (V, R_v)$  y  $M_H$  un matching máximo en  $H$  y  $U$  los vértices no saturados por  $M_H$ .  $|U| = n - 2|M_H|$  y  $|U| \leq |R_v \setminus M_H|$  (porque por cada arista que está en  $R_v \setminus M_H$  a lo sumo quedó un vértice en  $U$  porque  $M_H$  es máximo en  $H$ ) y  $|R_v \setminus M_H| = |R_v| - |M_H|$  porque  $M_H \subseteq R_v$ . Entonces  $n - 2|M_H| \leq |R_v| - |M_H|$ ,  $n \leq |R_v| + |M_H| \leq |R_v| + |M|$  (porque  $M_H$  es un matchig en  $G$  y  $M$  es máximo). ■

**Teorema 4.** Dado un grafo  $G$ , si  $I$  es un conjunto independiente máximo de  $G$  y  $R_n$  un recubrimiento mínimo de las aristas de  $G$ , entonces  $|I| + |R_n| = n$ .

*Demostración.* Se deriva del Lema 2 ■

## 5. Problema del cartero chino

Recordemos este problema.

**Definición 4.** Dado un grafo  $G = (V, X)$  con longitudes asignadas a sus aristas,  $l : X \rightarrow \mathbb{R}_{\geq 0}$ , el **problema del cartero chino** consiste en encontrar un circuito que pase por cada arista de  $G$  **al menos** una vez de longitud mínima.

Sabemos que si el grafo es euleriano, entonces la solución a este problema es un ciclo euleriano.

Si el grafo no es euleriano, vamos a tener que recorrer algunas aristas dos veces. Para determinar cuáles, podemos seguir el siguiente procedimiento:

1. Definimos  $S = \{v \in V : d(v) \text{ es impar}\}$ , el conjunto de los vértices que tienen grado impar. Llamemos  $s = |S|$  (seguro  $s$  es par).
2. Para cada par de estos vértices,  $v_i, v_j \in S$  calculamos,  $d(v_i, v_j)$ , la longitud del camino mínimo en  $G$ .
3. Armamos el grafo completo pesado,  $K_s$ , con los vértices de  $S$ , donde  $l(v_i, v_j) = d(v_i, v_j)$ .
4. Buscamos  $M$ , un matching perfecto (de cardinal  $s/2$ ) de peso mínimo en este grafo  $K_s$ .
5. Construimos el multigrafo  $G^*$  duplicando en  $G$  las aristas que definen un camino mínimo entre  $v_i$  y  $v_j$  para todas las aristas  $(v_i, v_j) \in M$ .  $G^*$  seguro es euleriano.
6. Buscamos un ciclo euleriano,  $Z$ , en  $G^*$ .
7.  $Z$  es un circuito óptimo para el problema del cartero chino sobre  $G$ .

El único paso que no sabemos hacer es el 4, pero puede realizarse en  $O(s^3)$ . Como el resto de los pasos también son polinomiales, ésto deriva en un algoritmo polinomial para el problema del cartero chino.

Si el grafo  $G = (V, X)$  es dirigido el procedimiento es similar. Sabemos que  $d_{IN}(v) = d_{OUT}(v)$  para todo  $v \in V$  si, y sólo si, el digrafo tiene un circuito euleriano. Si sucede eso, el circuito óptimo es un circuito euleriano y es computacionalmente fácil encontrarlo.





Si no sucede ésto, podemos armar un digrafo bipartito completo orientado de la siguiente forma. Para cada vértice  $v$  con  $d_{IN}(v) - d_{OUT}(v) = r(v) > 0$ , colocamos  $r(v)$  vértices en uno de los conjuntos de la partición, por ejemplo en  $V_1$ . Y para cada vértice  $u$  con  $d_{OUT}(u) - d_{IN}(u) = s(u) > 0$ , colocamos  $s(u)$  vértices en el otro conjunto,  $V_2$ . Agregamos todos los arcos de los vértices de  $V_1$  a los de  $V_2$ , con peso igual a un camino mínimo orientado en  $G$  entre los vértices que representan. En este digrafo buscamos un matching perfecto de peso mínimo,  $M$  (existe matching perfecto porque  $|V_1| = |V_2|$  y se conocen algoritmos polinomiales para hacerlo). Luego, armamos  $G^*$  duplicando los arcos de  $G$  que pertenecen a los caminos mínimos que definen los pesos de los arcos de  $M$ . Ahora  $G^*$  tendrá un ciclo euleriano orientado que será la solución óptima del problema del cartero chino sobre  $G$ , repitiendo en el circuito los arcos agregados en  $G^*$ .

En cambio, para grafos mixtos (con aristas y arcos) no se conocen algoritmos polinomiales.

## 6. Bibliografía recomendada

- Sección 12.4 de J. Gross and J. Yellen, *Graph theory and its applications*, CRC Press, 1999.
- Capítulos 5 y 7 de J. Bondy and U. Murty, *Graph theory with applications*, Macmillan, 5ta edición, 1982.

## Referencias

- [1] D. Jungnickel. *Graphs, Networks and Algorithms, Fourth Edition*. Algorithms and Computation in Mathematics 5, Springer, 2013.