

Trabajo Práctico 3

[7529/9506] Teoría de Algoritmos I
Segundo cuatrimestre de 2021

Grupo:	404
Repositorio:	github.com/lucashemmingsen/7529tp3
Entrega:	nº 1 (08/11/2021)

Integrantes del grupo 404

Padrón	Apellido, Nombre	Email
76187	Xxxxxxxxxx, Xxxxx Xxxxx	xxxxxxxxxxxx@fi.uba.ar
97529	Xxxxxxxxx, Xxxxx Xxxxxxx	xxxxxxxxxx@fi.uba.ar
102593	Xxxxxx, Xxxxx Xxxxx	xxxxxxx@fi.uba.ar
102649	Xxxxxx, XXXXXXXX Xxxxx	xxxxxxx@fi.uba.ar
106713	Xxxx Xxxxx, Xxxxx	xxxxx@fi.uba.ar

Índice

I. Introducción	2
I.1. Resumen	2
I.2. Lineamientos básicos	2
P1Parte 1: Una campaña publicitaria masiva pero mínima	3
a. Enunciado	3
b. Formato de los archivos	3
1. Vuelos: Propuesta	4
2. Vuelos: Reducción algorítmica	5
3. Vuelos: Optimalidad	5
4. Vuelos: Programa	6
4.1. Uso	6
5. Vuelos: Complejidad	6
P2Parte 2: Equipos de socorro	8
a. Enunciado	8
1. Socorro: NP-Completo	9
2. Set-dominante: NP-Completo	10
3. Tratabilidad	11
P3Parte 3: Un poco de teoría	12
a. Enunciado	12
1. Reducción polinomial	13
2. Importancia de NP-Completo	13
3. Ejercicios teóricos	14

I. Introducción

I.1. Resumen

El presente informe documenta el enunciado y la solución del tercer trabajo práctico de la materia Teoría de Algoritmos I. El mismo comprende el análisis de los problemas planteados, la implementación y comparación de complejidad del primero de ellos, y varios puntos teóricos.

I.2. Lineamientos básicos

- El trabajo se realizará en grupos de cinco personas.
- Se debe entregar el informe en formato pdf y código fuente en (.zip) en el aula virtual de la materia.
- El lenguaje de implementación es libre. Recomendamos utilizar C, C++ o Python. Sin embargo si se desea utilizar algún otro, se debe pactar con los docentes.
- Incluir en el informe los requisitos y procedimientos para su compilación y ejecución. La ausencia de esta información no permite probar el trabajo y deberá ser re-entregado con esta información.
- El informe debe presentar carátula con el nombre del grupo, datos de los integrantes y y fecha de entrega. Debe incluir número de hoja en cada página.
- En caso de re-entrega, entregar un apartado con las correcciones mencionadas

P1. Parte 1: Una campaña publicitaria masiva pero mínima

a. Enunciado

Una empresa de turismo que vende excursiones desea realizar una campaña publicitaria en diferentes vuelos comerciales con el objetivo de llegar a todos los viajeros que parten del país A y que se dirigen al país B. Estos viajeros utilizan diferentes rutas (algunos vuelos directos, otros armando sus propios recorridos intermedios). Se conoce para una semana determinada todos los vuelos entre los diferentes aeropuertos con sus diferentes capacidades. Además parten del supuesto que durante ese periodo la afluencia entre A y B no se verá disminuida por viajes entre otros destinos.

Desean determinar en qué trayectos simples (trayecto de un viaje que inicia desde un aeropuerto y termina en otro) poner publicidad de forma de alcanzar a TODAS las personas que tienen el destino inicial A y el destino final B. Pero además desean que siempre que sea posible seleccionen la combinación que tenga el menor número de vuelos comerciales. Esto es porque pagan tanto por cantidad de vuelos como por pasajeros que cumplan con la condición de ser del país de origen A y con destino final B.

Se pide:

1. Proponer una solución algorítmica que resuelva el problema de forma eficiente. Explicarla paso a paso. Utilice diagramas para representarla.
2. Plantear la solución como si fuese una reducción de problema. ¿Puede afirmar que corresponde a una reducción polinomial? Justificar.
3. ¿Podría asegurar que su solución es óptima?
4. Programe la solución
5. Compare la complejidad temporal y espacial de su solución programada con la teórica. ¿Es la misma o difiere?

b. Formato de los archivos

El programa debe recibir por parámetro el path del archivo donde se encuentran los vuelos disponibles entre pares de ciudades y la cantidad maxima de pasajeros en la misma.

El archivo debe ser de tipo texto y presentar por renglón, separados por coma el pais de origen el de destino y la capacidad del mismo. Las primeras dos lineas del archivo contiene el pais de origen y de destino respectivamente

Ejemplo: "vuelos.txt"

```
A
B
A,D,140
A,E,200
D,B,100
...
```

Debe resolver el problema y retornar por pantalla la solución. Debe mostrar por consola en en que vuelos poner la publicidad. Además imprimir cual es la cantidad maxima de pasajeros que puede ir de A a B.

1. Vuelos: Propuesta

Enunciado P1.1

Proponer una solución algorítmica que resuelva el problema de forma eficiente. Explicarla paso a paso. Utilice diagramas para representarla.

Para la solución al problema planteado se propone reducir el problema a un problema de corte de flujo y solucionarlo con el algoritmo de Edmonds-Karp.

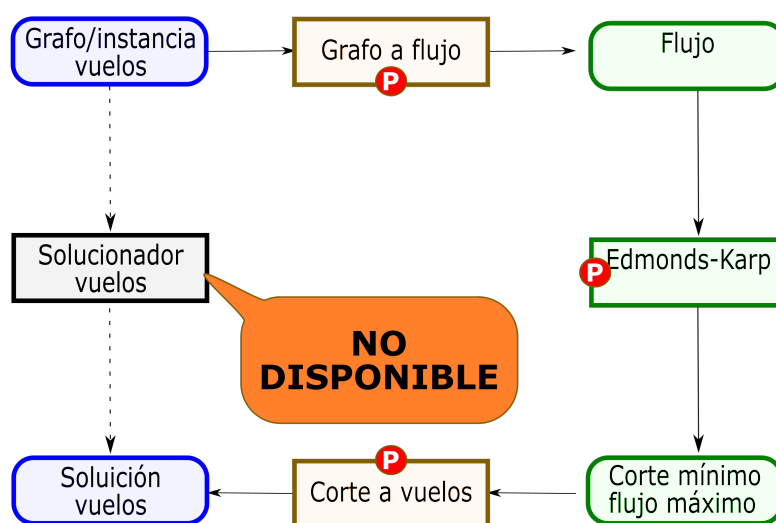


Figura 1: Reducción del problema, todas las operaciones son polinómicas.

En resumen, la solución completa sigue los siguientes pasos:

1. Procesar el archivo en el formato dado a un grafo dirigido, utilizando lista de adyacencias.
 - Cada nombre de aeropuerto es tratado como un número indicando la posición, en la lista de adyacencias, de cada lista de destino para un mismo origen.
 - Cada elemento está compuesto por una tupla de id de destino y "peso." asociado.
 - La estructura permite que peso pueda no ser simplemente un número, y el alias puede no ser un string. Pero en este punto, por el formato de entrada, sí lo serán.
2. Convertir el grafo en un diagrama de flujo, en el que:
 - En cada arco original se almacena:
 - una **capacidad máxima**, cuyo valor está dado por el anterior peso numérico del grafo.
 - y un **flujo**, inicialmente 0 (cero).
 - Cada arco original (a través de la clase `ArcoDirecto`) se obtiene un valor dado por la capacidad residual (la capacidad - el flujo). En el sentido contrario del grafo se almacena una referencia (a través de `ArcoInverso`) que, como valor, devuelve el flujo del arco directo (en el sentido original).

Por ejemplo, un arco $A \xrightarrow{10} B$ es reemplazado por un `ArcoDirecto` $A \xrightarrow{0/10} B$ que devuelve un valor $10 - 0 = 0$, y se agrega un nuevo arco $B \rightarrow A$ (que referencia al anterior) y devuelve un valor igual al flujo (0).

3. Se aplica Edmonds-Karp:

- a) Se inicializa en 0 un contador interno de flujo aumentado.
- b) Mientras haya un camino, encontrado con BFS, desde el origen y el destino dados en el archivo, cuyos arcos tengan valores mayores que cero:¹
 - 1) Busca el cuello de botella (el menor de todos los valores).
 - 2) Aumenta el camino: Si es un arco directo, aumenta el flujo (reduciendo su "valorz aumentando el del arco inverso asociado); si es un arco inverso, disminuye su flujo (aumentando su "valor", y reduciendo el del arco directo, ya que es el valor residual).
 - 3) Se aumenta un contador interno de flujo aumentado con el valor del cuello de botella.
4. Al terminar, se recorre con un BFS todos los nodos a los que pueda llegarse por un arco no saturado, es decir donde no haya un valor directo de 0 (si lo hay se lo agrega aun subconjunto de arcos **corte**), y se los agrega al subconjunto A; el resto de los nodos pertenecen al subconjunto B.
5. Se devuelve el contador de flujo como capacidad máxima de pasajeros del origen al destino, y el conjunto de arcos **corte** como los vuelos que cumplen con el criterio dado.

2. Vuelos: Reducción algorítmica

Enunciado P1.2

Plantear la solución como si fuese una reducción de problema. ¿Puede afirmar que corresponde a una reducción polinomial? Justificar

Puede considerarse una reducción algorítmica ya que se convirtió el grafo con las cantidades de pasajeros por vuelo, en un diagrama de flujo para resolverlo con Edmonds-Karp, y luego volvió a convertirse en un grafo (dado por los nodos A, B y los arcos de corte). Ambas transformaciones y la resolución son polinómicas, por lo que la **reducción es polinómicas**.

3. Vuelos: Optimalidad

Enunciado P1.3

¿Podría asegurar que su solución es óptima?

El algoritmo de Edmonds-Karp es una variante más específica de Ford Fulkerson,² por lo que también es óptimo para encontrar el flujo máximo y cuál es el conjunto de ArcoInverso por los que todo el flujo está pasando entre un subgrupo A que contiene a la fuente, y otro B que contiene al sumidero.

¹Esto es, para un arco directo significa que hay un valor residual, el flujo no llegó a su capacidad máxima. para un arco inverso significa que hay un flujo en el sentido contrario que podría disminuirse.

²especifica el criterio de selección del camino de aumento, mientras que Ford Fulkerson no.

4. Vuelos: Programa

Enunciado P1.4

Programa la solución

4.1. Uso

Para ejecutar el programa es necesario [descargar e instalar Python 3.10](#) y, desde la línea de comandos ubicado en la carpeta donde se descomprimió,³ ejecutar:

```
python -m src.vuelos NOMBRE_DE_ARCHIVO
```

Siendo NOMBRE_DE_ARCHIVO la ruta (relativa a la carpeta de trabajo) al archivo con las rutas, según el formato indicado en Formato de archivos. Por ejemplo, para usarse con uno de los ejemplos visto es clase:

```
python -m src.vuelos tests/entradas/test_diapo3.txt
```

newpage

5. Vuelos: Complejidad

Enunciado P1.5

Compare la complejidad temporal y espacial de su solución programada con la teórica. ¿Es la misma o difiere?

Para analizarla tomamos como referencia los pasos indicados en el punto P1.1.

1. Procesar el archivo puede asumirse una complejidad $O(E)$: consiste en insertar como mucho 2 sublistas en una lista (la de adyacencias) y un arco como elemento de una de esas sublistas, por cada arco (E). Espacialmente, es $O(V + E)$
2. **Conversión de grafo de flujo.** Se debe iterar cada arco realizando una inserción y una modificación (pueden presumirse $O(1)$ cada una). Por lo tanto, temporalmente sería $O(E)$. Si bien aumenta su tamaño, se duplican las aristas, por lo que espacialmente sigue siendo $O(V + E)$.
3. Edmonds-Karp tiene $O(VE)$ operaciones de empujapresumiblemente $O(E)$, donde:
 - 3.b Se aplica BFS, nuestra implementación de BFS es $O(V + E)$ temporalmente (iteramos dos veces por nodo V , y recorremos una vez cada arco E) y $O(V)$ espacialmente (V para nodo anterior, V para nodos a visitar y V para el camino).
 - 3.1 Cuello de botella, es una simple iteración por cada nodo del camino, almacenando el mínimo. $O(V)$ temporalmente, $O(1)$ espacialmente.
 - 3.2 Aumentar el camino, nuevamente una iteración por cada nodo del camino. $O(V)$ temporalmente.
4. **Computar subgrupos A, B y el corte.** Es un BFS, que aunque puede terminar antes podría ser que B sea el sumidero y por lo tanto seguiría siendo temporalmente $O(V + E)$. Espacialmente tenemos que $A + B = V$, y que $corte \leq E$, por lo tanto es $O(V + E)$.
- 5 **Convertir resultado:** se debe recorrer cada arco para convertirlo e imprimirlo. Es temporalmente $O(E)$, y como se necesita convertir uno por uno también lo es espacialmente.

En nuestra implementación el impacto en la complejidad temporal sigue dado por Edmonds-Karp, $O(VE^2)$, pudiendo descartarse las operaciones $O(V + E)$, $O(V)$ y $O(E)$. Con respecto **al problema original**, la complejidad temporal es $O(VE^2)$ siendo V la cantidad de aeropuertos y E la de viajes. Espacialmente, en el momento de mayor consumo hay 2 instancias de $O(V + E)$ por lo que pertenece a esa complejidad.

P2. Parte 2: Equipos de socorro

a. Enunciado

El sistema ferroviario de un país cubre un gran conjunto de su territorio. El mismo permite realizar diferentes viajes con transbordos entre distintos ramales y subramales que pasan por sus principales ciudades. Dentro de su proceso de mejoramiento del servicio buscan que ante una emergencia en una estación se pueda llegar de forma veloz y eficiente. Consideran que eso se lograría si el equipo de socorro se encuentra en esa misma estación o en el peor de los casos en una estación vecina (que tenga una trayecto directo que no requiere pasar por otras estaciones). Como los recursos son escasos desean establecer la menor cantidad de equipos posibles (un máximo de k equipos pueden solventar). Se solicita nuestra colaboración para dar con una respuesta a este problema.

Se pide:

1. Utilizar el problema conocido como “set dominante” para demostrar que corresponde a un problema NP-Completo.
2. Asimismo demostrar que el problema set dominante corresponde a un problema NP-Completo.
3. Con lo que demostró responda: ¿Es posible resolver de forma eficiente (de forma “tratable”) el problema planteado?

HINT: podría ser una buena idea utilizar 3SAT o VERTEX COVER.

1. Socorro: NP-Completo

Enunciado P2.1

Utilizar el problema conocido como “set dominante” para demostrar que corresponde a un problema NP-Completo.

El problema de situar los equipos de socorro revuelve en torno a el alcance de estos con respecto al conjunto de las estaciones. Se busca que k equipos de emergencia alcancen para que toda estación esté cubierta. Dado que un equipo sólo puede interactuar con su estación o las adyacentes a esta, el problema resulta ser una variante del problema “determinar set dominante”. En este caso, este problema está restringido a determinar un set dominante S con hasta k elementos.

«Un "set dominante" S de un grafo G es un subconjunto de vértices de dicho grafo, tales que cada vértice del grafo que no se encuentre en el subconjunto S , es adyacente a aquellos que sí.» ⁴

El algoritmo solucionador, a grosso modo, debe determinar un set dominante S , de límite de tamaño k , el cual resulte ser el posicionamiento preferible de los equipos de socorro. El problema de «hallar set dominante» fue probado NP-Completo (consecuencia de que Richard M. Karp determinó que «set cover problem» es NP-Completo, ocasionada debido a que que toda instancia de dominating set [“hallar set dominante”] es reducible a una de set cover problem y viceversa). Por lo tanto, y siendo que el algoritmo es básicamente una solución del problema “hallar set dominante”, el problema de ubicar los equipos de socorro es NP-Completo.

⁴Wayne Goddard, Michael A. Henning. «Independent domination in graphs: A survey and recent results». Discrete Mathematics, Volume 313, Issue 7, 2013, pág. 839-854, ISSN 0012-365X, <https://doi.org/10.1016/j.disc.2012.11.031>

2. Set-dominante: NP-Completo

Enunciado P2.2

Asimismo demostrar que el problema set dominante corresponde a un problema NP-Completo.

Set dominante es NP: Dado un set puedo recorrer cada vértice, marcarlo con sus adyacentes y comprobar que todos los vértices del grafo fueron marcados. Por lo tanto set dominante $\subseteq NP - Hard$.

3. Tratabilidad

Enunciado P2.2

Con lo que demostró responda: ¿Es posible resolver de forma eficiente (de forma “tratable”) el problema planteado?

No se ha llegado a completar este punto.

P3. Parte 3: Un poco de teoría

a. Enunciado

1. Defina y explique qué es una reducción polinomial y para qué se utiliza.
2. Explique detalladamente la importancia teórica de los problemas NP-Completos.
3. Tenemos un problema A, un problema B y una caja negra NA y NB que resuelven el problema A y B respectivamente. Sabiendo que B es P.
 - a) Qué podemos decir de A si utilizamos NA para resolver el problema B (asumimos que la reducción realizada para adaptar el problema B al problema A es polinomial)
 - b) Qué podemos decir de A si utilizamos NB para resolver el problema A (asumimos que la reducción realizada para adaptar el problema A al problema B es polinomial)
 - c) ¿Qué pasa con los puntos anteriores si no conocemos la complejidad de B, pero sabemos que A es NP-C?

1. Reducción polinomial

Enunciado P3.1

Defina y explique qué es una reducción polinomial y para qué se utiliza.

Una reducción polinomial es un procedimiento por el cual se transforma una instancia de un problema cuya solución no es conocida en otra instancia de un problema el cual si se sabe resolver. Luego se transforma la solución obtenida para el problema conocido en una solución para el problema original.

Si las transformaciones se realizan en tiempo polinomial, entonces es una reducción polinomial.

Notación: Dados los problemas X e Y , se dice $Y \leq_p X$ si Y es polinomialmente reducible a X . La reducción polinomial puede utilizarse como «caja negra» para resolver problemas cuya solución es desconocida. También como medida de complejidad en el caso de saber la complejidad de alguno de los problemas. Si $Y \leq_p X$, entonces podemos decir que X es «al menos tan difícil de resolver» que Y .

2. Importancia de NP-Completo

Enunciado P3.2

Explique detalladamente la importancia teórica de los problemas NP-Completo.

Un problema es NP-Completo si pertenece a NP y también pertenece NP-Hard, lo cual quiere decir que es uno de los problemas más difíciles de resolver dentro de NP. Esta condición de ser de los problemas más difíciles de resolver dentro de NP los hace importantes ya que si se encontrara la solución a un problema NP-Completo en tiempo polinómico, se probaría que $P = NP$.

Esto ocurre dado que sabemos que $P \subseteq NP$,⁵ y si se pudiera probar que el problema más difícil de NP puede solucionarse en tiempo polinomial, entonces todos los problemas de NP se podrían solucionar en tiempo polinomial y por lo tanto $NP \subseteq P$.

Finalmente, si $NP \subseteq P$ y $P \subseteq NP$, entonces $P = NP$.

⁵Si se tiene un algoritmo P que resuelve un problema de decisión, puede usarse como "algoritmo certificador" NP, para verificar la solución dada sea correcta simplemente comparando que sean la misma. Por esto, todo algoritmo de decisión P es también NP; pero no necesariamente al revés.

3. Ejercicios teóricos

Enunciado P3.2

Tenemos un problema A, un problema B y una caja negra NA y NB que resuelven el problema A y B respectivamente. Sabiendo que B es P.

1. Qué podemos decir de A si utilizamos NA para resolver el problema B (asumimos que la reducción realizada para adaptar el problema B al problema A es polinomial)
2. Qué podemos decir de A si utilizamos NB para resolver el problema A (asumimos que la reducción realizada para adaptar el problema A al problema B es polinomial)
3. ¿Qué pasa con los puntos anteriores si no conocemos la complejidad de B, pero sabemos que A es NP-C?

1. Qué podemos decir de A si utilizamos NA para resolver el problema B (asumimos que la reducción realizada para adaptar el problema B al problema A es polinomial)

$B \leq_p A$. Que A es al menos tan difícil de resolver como B. No me dice nada sobre la complejidad de A.

2. Qué podemos decir de A si utilizamos NB para resolver el problema A (asumimos que la reducción realizada para adaptar el problema A al problema B es polinomial)

$A \leq_p B$. Se puede decir que $A \subseteq P$. Si B es igual o más difícil de resolver que A, y B es P; entonces A tiene que ser P.

3. ¿Qué pasa con los puntos anteriores si no conocemos la complejidad de B, pero sabemos que A es NP-C?

A es NP-C \implies A es NP-H y A es NP.

- 1) $B \leq_p A$ Puedo decir que B es al menos NP.
- 2) $A \leq_p B$ Entonces puedo decir que B es NP-Hard.