



DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Práctico I

Algoritmos y Estructuras de Datos III
Primer Cuatrimestre de 2015

Integrante	LU	Correo electrónico
Aldasoro Agustina	86/13	agusalaldasoro@gmail.com
Noriega Francisco	XXX/XX	mail
Zimenspitz Ezequiel	155/13	ezeqzim@gmail.com
Zuker Brian	XXX/XX	mail



Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Ciudad Universitaria - (Pabellón I/Planta Baja)

Intendente Güiraldes 2160 - C1428EGA

Ciudad Autónoma de Buenos Aires - Rep. Argentina

Tel/Fax: (54 11) 4576-3359

<http://www.fcen.uba.ar>

Resumen

Habiéndonos sido dado una serie de tres problemáticas a resolver, se plantean sus respectivas soluciones acorde a los requisitos pedidos. Se adjunta una descripción de cada problema y su solución, conjunto a su análisis de correctitud y de complejidad sumado a su experimentación. El lenguaje elegido para llevar a cabo el trabajo es C++.

Índice

1. Problema 1: ZombieLand	3
1.1. Descripción de la problemática	3
1.2. Resolución propuesta y justificación	4
1.3. Análisis de la complejidad	5
1.4. Código fuente	6
1.5. Experimentación	7
1.5.1. Contrastación Empírica de la complejidad	7
1.5.2. Comparar países de a dos	8
1.5.3. Modificación en el algoritmo	9
2. Problema 2: Alta Frecuencia	10
2.1. Descripción de la problemática	10
2.2. Resolución propuesta y justificación	11
2.3. Análisis de la complejidad	12
2.4. Código fuente	13
2.5. Experimentación	14
2.5.1. Contrastación Empírica de la complejidad	14
3. Problema 3: El señor de los caballos	15
3.1. Descripción de la problemática	15
3.2. Resolución propuesta y justificación	17
3.3. Análisis de la complejidad	18
3.4. Código fuente	19
3.5. Experimentación	20
3.5.1. Contrastación Empírica de la complejidad	20

1. Problema 1: ZombieLand

1.1. Descripción de la problemática

Dado un mapa con n ciudades, en cada una de ellas se encuentra una determinada cantidad de Zombies y una determinada cantidad de Soldados (ambas también dadas). El objetivo del problema es exterminar a la invasión zombie, para ello es necesario un enfrentamiento *zombies vs soldados* por cada ciudad. Con el fin de que este enfrentamiento sea positivo, es decir se logre matar a todos los zombies de una ciudad, es necesario que la cantidad de zombies sea, a lo sumo, diez veces más grande que la cantidad de soldados.

Como los soldados se encuentran atrincherados, se puede optar entre llevar a cabo un enfrentamiento o no (ya que no se desea provocar un ataque donde ya se sabe que se va a perder). Los soldados acuartelados no pueden moverse de la ciudad en la que están, pero sí se cuenta con una dotación de soldados extra que se la puede ubicar en cualquiera de las n ciudades acorde a lo deseado. La cantidad de soldados extra es ilimitada, mas los recursos para trasladarlos no lo son. El traslado de cada soldado extra a una ciudad tiene un costo específico que varía acorde a la ciudad elegida. Siempre que sea económicamente posible, se puede trasladar una cantidad de soldados indistinta por ciudad.

Debido a que los recursos económicos son finitos, no siempre va a ser posible salvar a las n ciudades. Lo que se desea en este problema es maximizar la cantidad de ciudades salvadas, respetando el presupuesto. Es decir, se deben establecer las cantidades de soldados extras enviados a cada ciudad de modo que la cantidad de ciudades salvadas sea la óptima y gastando un monto por debajo del presupuesto. El algoritmo debe tener una complejidad temporal de $O(n \cdot \log(n))$, siendo n la cantidad de ciudades del país.

Aca se podría poner unos dibujitos de soluciones óptimas como para que quede más lindo

1.2. Resolución propuesta y justificación

Ver que habla mucho sobre vectores y dijo que no mezcláramos implementación con algoritmo

Para la resolución del problema decidimos utilizar un algoritmo goloso, el mismo tomará una decisión óptima para cada instante, sin revisar si será la mejor a nivel global.

Como primera instancia, el algoritmo simplemente calcula, para cada ciudad, cuánto sería el costo de salvarla. Para ello, primero se calcula la cantidad de soldados extra necesarios y luego se multiplica por el costo de traslado de cada unidad:

```
soldados_extras_necesarios = redondeo_hacia_arriba((zombies - (soldados_existentes * 10)) / 10)
costo_total = costo_unitario * soldados_extras_necesarios
```

Luego de haber obtenido una magnitud con la cual se pueden comparar las ciudades entre sí, se ordenan las ciudades de menor a mayor en base al costo de salvarla para ser recorridas secuencialmente y enviar los ejércitos requeridos hasta que alcance el presupuesto.

Como las ciudades van a estar almacenadas en un vector, en la posición vec_0 se va a encontrar la ciudad que es más barata de salvar y en la posición vec_{n-1} la más cara. En este vector se van a encontrar todas las ciudades pasadas por parámetro, sin importar que una ciudad ya este salvada al ser ingresada.

Se recorre este vector secuencialmente, de modo que para cada posición $i \in [0, n-1]$ se va a comparar el costo de salvar la ciudad i contra el presupuesto restante en esa instancia (`presupuesto_actual`). Si es una compra factible, se resta el `costo_total` del `presupuesto_actual` y se envían las tropas necesarias a la ciudad i ; en caso contrario se le asigna 0 a las tropas extras enviadas a la ciudad i .

Por la forma en que fue construido el vector, cuando se encuentre una ciudad i que el costo de salvarla excede del `presupuesto_actual` las demás ciudades $j \in [i+1, n-1]$ tampoco podrán ser salvadas. De todos modos, nuestro algoritmo realiza el mismo cálculo para todas las posiciones del vector.

Como en este vector las ciudades no aparecen en orden creciente por su número, debemos reordenarlas. Debido a que el *id* de cada ciudad va a ser único y va a encontrarse en el intervalo $[0, n-1]$, se lo recorre secuencialmente colocando cada ciudad en la posición de su *id* dentro de un nuevo vector *respuesta*. Ver que este párrafo me aparece que no quedo muy claro...

El algoritmo resuelve el problema salvando la mayor cantidad de ciudades posibles porque **INSERT FORMAL DEMO**

1.3. Análisis de la complejidad

La complejidad de nuestra solución es $O(n \cdot \log(n))$, siendo n la cantidad de ciudades del país.

Como primera instancia, el algoritmo almacena las ciudades con sus características (pasadas como parámetro) dentro de un vector para luego poder utilizarlas de un modo práctico. Como esto lo realiza secuencialmente, tiene un costo lineal $O(n)$. **Esto habia que ponerlo aca asi? o directamente nada? o solo nombrarlo?**

Algorithm 1: zombieland

```
for cada ciudad en vector país do
    Calcular la cantidad de soldados extra necesarios y el costo de esto.
    Almacenar esta información en un vector datos mediante push_back()
Ordena al vector datos mediante sort_heap()
for cada ciudad en vector datos do
    Calcular si puede ser salvada.
    Indicarle cantidad de soldados extras enviados y cambios en el presupuesto_actual.
for cada ciudad en vector datos do
    Insertar en el vector respuesta[ciudad.id] la ciudad actual.
```

En el codigo, yo pondria este reordenamiento dentro de zombieland, ya que hay que considerarlo para medir tiempos.

El código presenta un **primer ciclo for** que calcula el costo de salvar a cada ciudad y las agrega mediante `push_back()` a un vector. El ciclo recorre linealmente todas las ciudades por lo que tiene complejidad $O(n)$.

El cálculo de salvar a cada ciudad coincide con el descrito en la sección anterior, el cual por ser operaciones aritméticas es $O(1)$. Armar el nuevo struct para insertar dentro del vector `datos` también posee un costo constante $O(1)$. La función `push_back()`¹ tiene costo $O(1)$ amortizado, lo que implica que cuando no precisa redimensionar el vector cuesta esto, y cuando lo hace, toma tiempo lineal en la cantidad de elementos. Como insertamos durante todo el ciclo tomamos el costo amortizado $O(1)$.

Por lo tanto, la complejidad total del primer ciclo for nos da $O(n)$.

Le sigue **ordenar el vector** con estos datos, para ello usamos `sort_heap()`² de la librería de C cuya complejidad es $O(n \cdot \log(n))$.

A continuación, se realiza un **último ciclo for** que salva las ciudades que pueda, mientras dure el presupuesto y deja en 0 soldados enviados a las ciudades que no pueden ser salvadas. Estos cálculos aritméticos y asignaciones son todos de complejidad constante $O(1)$. Este ciclo recorre linealmente todas las ciudades por lo tanto, lo hace con complejidad $O(n)$.

Finalmente, se debe **reordenar el vector** obtenido hasta ahora para que quede en orden creciente respecto de su *id*. Como esto se hace recorriendo secuencialmente el primer vector, asignándole uno a uno los elementos al nuevo vector `respuesta` indexados; sólo hace una pasada lineal con costo $O(n)$.

Como cada paso de los mencionados son secuenciales, las complejidades se suman, obteniendo: $O(n) + O(n \cdot \log(n)) + O(n) + O(n)$ que es igual a $O(n \cdot \log(n))$ por propiedades de O .

¹http://www.cplusplus.com/reference/vector/vector/push_back/

²http://www.cplusplus.com/reference/algorithm/sort_heap/?kw=sort_heap

1.4. Código fuente

```
struct ciudad{
    int zombies;
    int soldados;
    int costo;
};
```

```
struct ciudad2{
    int numCiudad;
    int soldadosNecesarios;
    int costoTotal;
    bool operator< (const ciudad2& otro) const{
        return costoTotal < otro.costoTotal;
    }
};
```

Poner como leemos y escribimos?? y el main? CREO QUE SI...

Algorithm 2: ZombieLand(out: vector<ciudad2>; in: int cantCiudades, int presupuesto, vector<ciudad>& pais; in/out: int& salvadas)

```
salvadas = 0;
vector<ciudad2> datos;
for (int i = 0; i < cantCiudades; ++i) do
    ciudad2 actual;
    actual.numCiudad = i;
    double diferencia = (pais[i].zombies - pais[i].soldados * 10);
    if (diferencia > 0) then
        | actual.soldadosNecesarios = ceil(diferencia/10);
    else
        | actual.soldadosNecesarios = 0;
    actual.costoTotal = actual.soldadosNecesarios * pais[i].costo;
    datos.push_back(actual);
sort_heap(datos.begin(), datos.end());
for (int i = 0; i < cantCiudades; ++i) do
    int dif = presupuesto - datos[i].costoTotal;
    if (dif >= 0) then
        | salvadas++;
        | presupuesto = dif;
    else
        | datos[i].soldadosNecesarios = 0;
return datos;
```

No seguí poniendo el algoritmo, por si le tenemos que hacer algún cambio.

1.5. Experimentación

Ver bien como generar ciudades aleatorias

1.5.1. Contrastación Empírica de la complejidad

-Comparar tiempos para ciudades de tamaño 10, 100, 1000, 10000 y 100000 generando aleatoriamente varias de cada una y sacar promedio. Hacer lo que hicieron en clase

1.5.2. Comparar países de a dos

Consideramos 10?? casos, donde para cada uno tomamos dos países (A y B) con la misma cantidad de ciudades **para cada par de países, cantidades distintas de ciudades (100, 500, 1000, ... (hasta donde nos de la compu jajaj))**. Estos casos los vamos a armar de tal forma que A tenga ya todas sus ciudades salvadas al momento de ser ingresada, en cambio B tiene una relación de cantidad soldados/zombies aleatoria. Corremos el algoritmo para todos los pares de países.

Dado a que nuestro algoritmo no establece ningún control de filtro sobre casos donde las ciudades ya esten salvadas al momento de ser recibidas como parámetro, consideramos que para cada par de ciudades los tiempos de cómputo no van a diferir mucho.

Aca va el gráfico donde muestra la diferencia entre cada par. Y ver si lo dibujamos linealizado como mostraron en clase tambien.

Aca hay que escribir si coincidió con lo que creiamos o no y dar una pequeña explicación de porque pasa eso

1.5.3. Modificación en el algoritmo

Decidimos hacer una modificación dentro del algoritmo. Creemos que esta misma va a generar una mejora en la cantidad de tiempo.

La misma consiste en: modificar el último *ciclo for* para que cuando encuentre la primer ciudad que no pueda salvar deje de iterar ya que las restantes tampoco podrán ser salvadas (esto ocurre porque están ordenadas por costo).

Consideramos que saliendo antes del ciclo, el tiempo de cómputo para un mismo país debería ser menor que con el algoritmo original.

Experimentamos 10 casos distintos de tamaño 1000(?) y para cada uno comparamos sus tiempos de ejecución entre el algoritmo original y el modificado.

Aca va el gráfico de eso que acabo de explicar arriba.

2. Problema 2: Alta Frecuencia

2.1. Descripción de la problemática

Se quiere transmitir información secuencialmente mediante un enlace el mayor tiempo posible. Los enlaces tienen asociadas distintas frecuencias, con un costo por minuto y un intervalo de tiempo (sin cortes) en el cual funcionan. Se utilizan durante minutos enteros, y es posible cambiar de una frecuencia a otra instantáneamente (del minuto 1 al 4 uso la frecuencia A y del 4 al 6 la B). Los datos del precio y e intervalo de tiempo de cada frecuencia son dados. Se desea optimizar este problema para transmitir todo el tiempo que tenga al menos una frecuencia abierta, pero gastando la menor cantidad de dinero. Se debe contar con una complejidad de $O(n \cdot \log(n))$.

A continuación se muestran dos casos particulares de este problema. En ambos se ofrecen tres frecuencias, con distintos costos cada una. Se puede ver recuadrado en violeta cuál es la elección que debe hacerse por intervalo de tiempo.

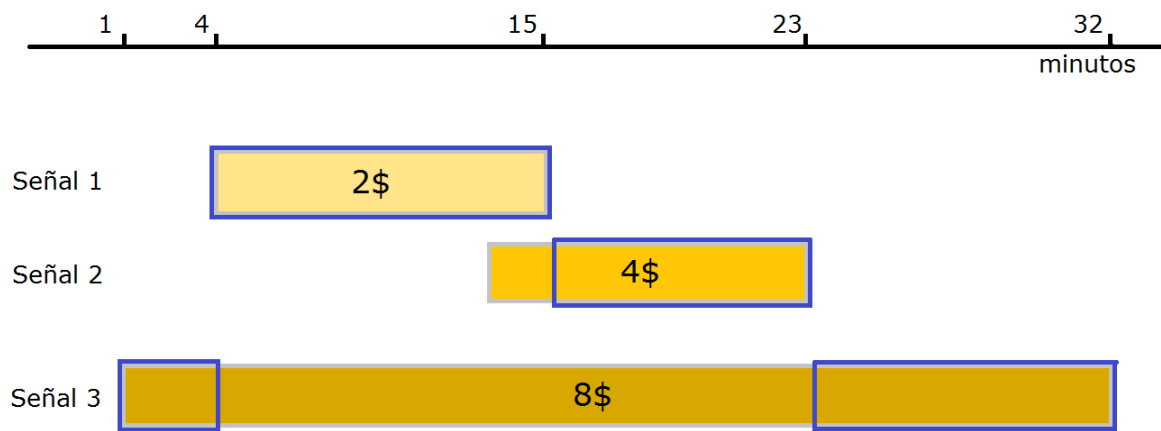


Figura 1: Ejemplo 1

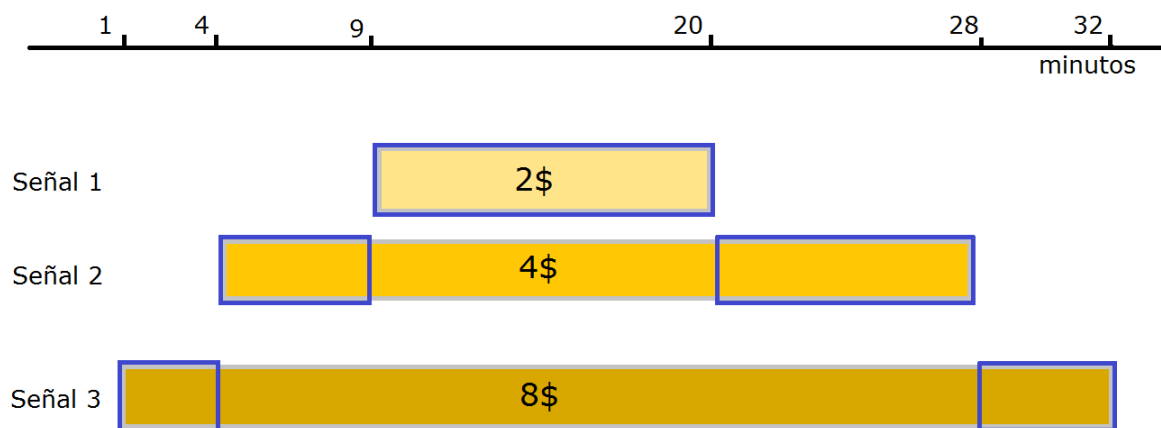


Figura 2: Ejemplo 2

2.2. Resolución propuesta y justificación

El algoritmo que utilizamos pertenece a la familia de *Divide & Conquer*.

Aca pasa lo mismo de la implementacion porque hablamos mucho de vector y bla...

Todas las frecuencias se encuentran almacenadas en un vector. Primero se las ordena en orden creciente respecto del costo, es decir en la primera posición se almacena la más barata y en la última, la más cara. Una vez que contamos con este ordenamiento inicial, se va a proseguir mediante el *divide* dentro de este vector.

Se prosigue de acuerdo a las características de *Divide & Conquer*. De modo recursivo, se divide al vector pasado por parámetro por la mitad creando dos nuevos a los cuales se les aplica nuestro algoritmo de *merge*. El caso base se da cuando el vector tiene un sólo elemento, para lo cual se devuelve el vector tal cual entró.

Nuestro algoritmo de *Merge* se encarga de elegir entre las frecuencias de dos arreglos pasados como parámetro, de modo que devuelve un sólo vector indicando los intervalos ocupados por las frecuencias elegidas.

Dado el enunciado del problema, para elegir qué frecuencia utilizar en determinado intervalo de tiempo se debe priorizar el precio más barato emitiendo señal siempre que sea posible. Gracias a que en el primer llamado de nuestra función estas se encontraban en orden creciente respecto del costo, en cada paso de *merge* de los dos arreglos de entrada con intervalos se van a priorizar los de la izquierda. Es decir que en cada paso, todos los intervalos pertenecientes al vector de la izquierda (como son los de menor precio) van a pertenecer al vector resultado. Mientras que sólo los intervalos que completen tiempo sin señal pertenecientes al vector de la derecha van a ser colocados en el vector resultado. **Que alguien lea esto para ver si se entiende, para mi si :)**

Una vez terminado este paso recursivo, vamos a contar con el conjunto de intervalos que completen la mayor cantidad de tiempo con el costo más bajo. Si no hay dos frecuencias con el mismo costo, esta solución (óptima) es única, en caso contrario podría existir más de una solución óptima.

Este algoritmo resuelve lo propuesto dando una solución óptima porque...

2.3. Análisis de la complejidad

Primero consideramos que nos encontramos en un caso donde la solución óptima es única. Si contamos con la oferta de n frecuencias, podemos asegurar que la cantidad de intervalos que va a contener la salida es de, a lo sumo, $2n-1$. Esta cota superior está dada porque, si consideramos agregar de a una las distintas frecuencias (empezando por las de menor costo), lo máximo que puede agregar son dos intervalos (o cubrir huecos de otros que no llegaron a llenar dos). **Aca viene Eze y explica bien todo esto :D.**

Por lo tanto, al hacer nuestro algoritmo de Divide & Conquer vamos a contar con, a lo sumo, $2n-1$ intervalos. Esto nos otorga una complejidad de $O(n \cdot \log(n))$ por el Teorema de ??????.

2.4. Código fuente

2.5. Experimentación

2.5.1. Contrastación Empírica de la complejidad

-Comparar tiempos para 10, 100, 1000, 10000 y 100000 intervalos generando aleatoriamente varios de cada una y sacar promedio. Hacer lo que hicieron en clase

3. Problema 3: El señor de los caballos

3.1. Descripción de la problemática

En este problema, se presenta un tablero de ajedrez de tamaño $n \times n$, el cual cuenta con alguna cantidad de caballos ubicados en una posición aleatoria del tablero. Lo que se quiere lograr es *cubrir* todo el tablero. Un casillero se considera cubierto si hay un caballo en él o bien, si es una posición en la cual algún caballo existente puede moverse con un sólo movimiento. Para lograr este cometido, puede ser necesario agregar nuevas fichas *caballo* al tablero. No existe un límite en la cantidad de caballos para agregar, pero lo que se busca es dar una solución con la mínima cantidad de caballos posibles.

En la figura 3 se pueden ver todas las casillas que están cubiertas por un sólo caballo.

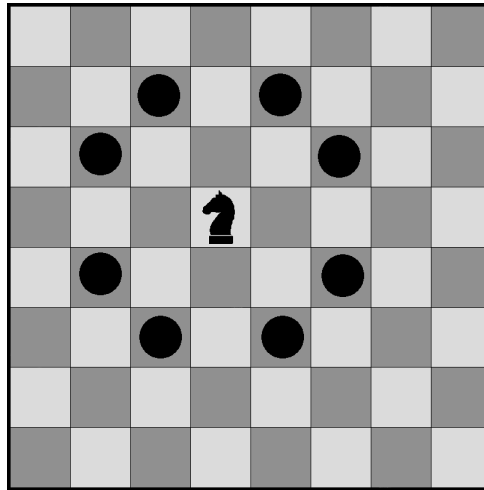


Figura 3: Casillas que *cubre* un caballo

3.2. Resolución propuesta y justificación

3.3. Análisis de la complejidad

3.4. Código fuente

3.5. Experimentación

3.5.1. Contrastación Empírica de la complejidad

-Hacer lo que hicieron en clase