



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	agusaldasoro@gmail.com
Noriega Francisco	660/12	frannoriega.92@gmail.com
Zimenspitz Ezequiel	155/13	ezeqzim@gmail.com
Zuker Brian	441/13	brianzuker@gmail.com



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++.

Estos son los comandos para compilar cada ejercicio (el flag se utilizó para la librería `chrono` para medir tiempos de ejecución):

```
g++ -o main Zombieland.cpp -std=c++11
```

```
g++ -o main AltaFrecuencia.cpp -std=c++11
```

```
g++ -o main SenorCaballos.cpp -std=c++11
```

Dentro de cada .cpp está el comando para compilar cada ejercicio desde la carpeta donde se encuentran los mismos.

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	7
1.5. Experimentación	8
1.5.1. Contrastación Empírica de la complejidad	8
1.5.2. Modificación del algoritmo	8
1.5.3. Comparar países de a dos	8
2. Problema 2: Alta Frecuencia	9
2.1. Descripción de la problemática	9
2.2. Resolución propuesta y justificación	10
2.3. Análisis de la complejidad	11
2.4. Código fuente	12
2.5. Experimentación	13
2.5.1. Contrastación Empírica de la complejidad	13
2.5.2. Modificación del algoritmo	14
3. Problema 3: El señor de los caballos	15
3.1. Descripción de la problemática	15
3.2. Resolución propuesta y justificación	16
3.3. Análisis de la complejidad	17
3.4. Código fuente	18
3.5. Experimentación	19
3.5.1. Contrastación Empírica de la complejidad	19

1. Problema 1: ZombieLand

1.1. Descripción de la problemática

En un país con n ciudades, se encuentran una determinada cantidad de Zombies y de Soldados por cada una de ellas. El objetivo del problema es exterminar la invasión zombie, para ello es necesario un enfrentamiento *zombies vs soldados* por cada ciudad. Para que el combate sea positivo en una ciudad, es decir se logre matar a todos los zombies de la misma, es necesario que la cantidad de zombies sea, a lo sumo, diez veces más grande que la cantidad de soldados.

Se sabe de antemano cuántos zombies y cuántos soldados se encuentran atrincherados en cada ciudad (o sea que cada ciudad sabe cuando iniciar el combate para no ser infectados). 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 para salvarla. La cantidad de soldados extra es ilimitada, mas los recursos para trasladarlos no lo son. El costo del traslado depende de cada ciudad. Siempre que se respete el presupuesto del país, se pueden trasladar todos los soldados necesarios para salvar a cada 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

Para la resolución del problema decidimos utilizar un algoritmo goloso, que salvará en cada paso a la ciudad que más le convenga en ese momento, es decir, la que permita maximar la cantidad de ciudades salvadas.

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 se agote el presupuesto.

Notar que si alguna ciudad no requiere soldados extras para ser salvada, entonces serán las primeras en ser salvadas dado que el costo_total será igual a 0.

Se recorren secuencialmente las ciudades ordenadas por el costo_total, de modo que para cada una se va a comparar el costo de salvarla contra el presupuesto restante en ese momento (presupuesto_actual). Si es factible el salvataje, se resta el costo_total del presupuesto_actual y se envían las tropas necesarias a la ciudad; en caso contrario se la marca como ciudad perdida.

Vale aclarar que el orden impuesto a las ciudades implica que cuando no se pueda salvar a una ciudad, no se puede salvar a ninguna otra.

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 parrafo me aparece que no quedo muy claro... es necesario este párrafo? habla del stdout... no me parece necesario en absoluto... yo lo sacaría*

El algortimo resuelve el problema salvando la mayor cantidad de ciudades posibles porque **FRAN NORIEGA – INSERT FORMAL DEMO**

1.3. Análisis de la complejidad

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

En primera instancia, guardamos los datos de las ciudades pasadas por stdin en structs y los dejamos dentro de un vector, para luego poder utilizarlas de un modo práctico. Como esto se realiza secuencialmente, tiene costo lineal $O(n)$.

Algorithm 1: zombieland

```

for each ciudad  $\in$  país do
  | Calcular la cantidad de soldados extra necesarios y el costo de salvarla.
  | Almacenar esta información en un vector datos mediante push.back()
Ordena al vector datos mediante sort()
while pueda salvar do
  | if puede ser salvada then
  | | Indicar cantidad de soldados extras enviados y actualizar el presupuesto_actual
es necesario este cacho? habla del stdout... no me parece necesario en absoluto... yo lo sacaría
for each ciudad en vector datos do
  | Insertar en el vector respuesta[ciudad.id] la ciudad actual.
  
```

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

A mi no me parece que sea necesario para medir tiempos... o sea, arrancar de la instancia pasada por parametro, salvo que la limemos y usemos avls, heaps, etc. no me parece necesario contarlas. lo mismo para escribir, la respuesta está, en el ej2 tenemos que calcular un pedazo de respuesta post algoritmo y lo hacemos y lo consideramos, el std out extra porque deberíamos considerarlo? es algo que nos piden para visualizar, si las cosas andan mal, porque pueden llegar a estar mal

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 descripto 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()*² cuya complejidad es $O(n \log(n))$.

A continuación, se realizan dos **último ciclos while** el primero salva todas las ciudades que pueda, mientras dure el presupuesto y el segundo deja en O soldados enviados a las ciudades que no pueden ser salvadas. Estos cálculos aritméticos y asignaciones son todos de complejidad constante $O(1)$. El primer ciclo toma $O(\text{ciudades_salvadas})$ y el segundo $O(n - \text{ciudades_salvadas})$ dando como resultado un recorrido lineal sobre todas las ciudades, por lo tanto lo hace con complejidad $O(n)$.

Para mi no va...

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 $O(n)$.

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

²<http://www.cplusplus.com/reference/algorithm/sort/>

Como cada paso de los mencionados son secuenciales, las complejidades se suman, obteniendo:

$O(n) + O(n \cdot \log(n)) + O(n)$ que es igual a **$O(n \cdot \log(n))$** por propiedades de O .

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... no se, se puede dejar como apendice esto no? directamente lo imprimimos del sublime, toda la paja pasar todos los algoritmos a latex, con el main y las funciones extras y bla bla bla

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. Constrastación 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

QUE ES ESTO??? MODIFIQUE EL ALGORITMO PARA QUE HAGA EXACTAMENTE ESTO ANTES DE LEERLO... ESTO NO ES UNA EXPERIMENTACIOOOOON, es una chamuyada asquerosa y sin sentido!

1.5.2. Modificación del 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.

1.5.3. 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

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.

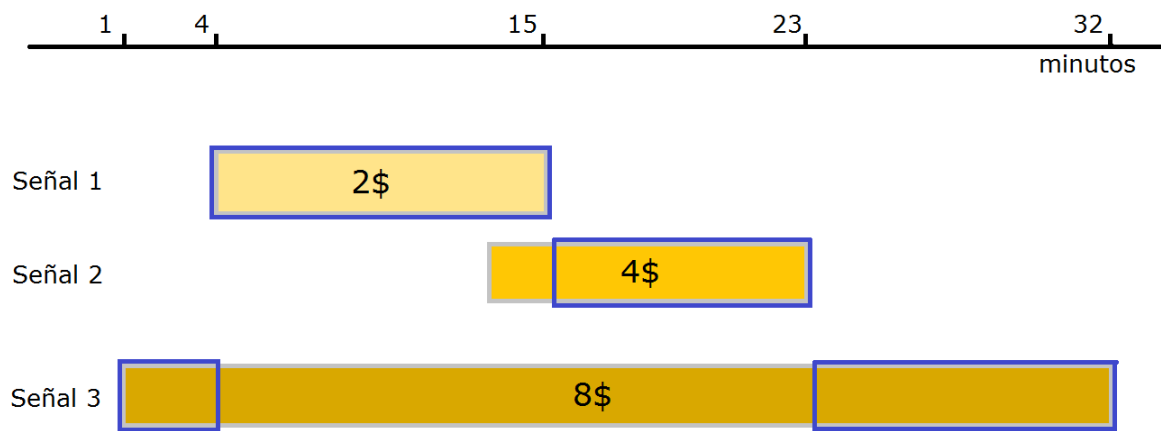


Figura 1: Ejemplo 1

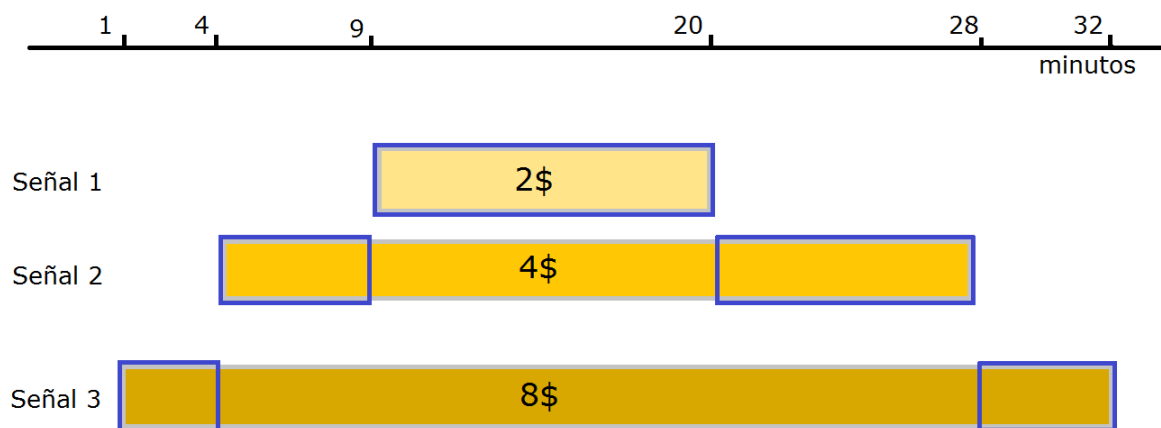


Figura 2: Ejemplo 2

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...

El primer paso consiste en ordenar las frecuencias de menor a mayor en base al costo de cada una.

Luego, se sigue el esquema clásico de Divide & Conquer:

```
divide(conjuntoDeFrecuencias F){  
    Si hay mas de un elemento:  
        Divido F en mitades A, B.  
        intervalosA = divide(A)  
        intervalosB = divide(B)  
        Devuelvo conquer(intervalosA, intervalosB)  
    Si hay un solo elemento:  
        Lo devuelvo.  
}
```

En palabras, si hay una sola frecuencia, la devolvemos, pues es trivial que su intervalo de duración es el más barato y el de mayor extensión temporal.

Si no, intervalosA será el conjunto de intervalos en la que funcionará cada frecuencia, de modo que el costo de contratar el servicio con este cronograma sea mínimo en el costo y máximo en la cantidad de tiempo de uso. E intervalosB será un conjunto con las mismas características, con la diferencia de que el A será el más óptimo de la mitad más barata y el B el más óptimo de la mitad más cara. Esto resulta de haber ordenado las frecuencias por su costo antes de comenzar con este tramo de algoritmo.

Al hacer `conquer(intervalosA, intervalosB)` se obtiene el conjunto de intervalos con las características mencionadas pero de todas las frecuencias.

Este último paso abusa del invariante: todos los intervalos del conjunto intervalosA deben aparecer en el conjunto solución, y que lo único que debe agregar son los intervalos de intervalosB que o bien aumentan el rango de tiempo para transmitir (uso el servicio desde antes o más tiempo) o bien completan gaps que puedan existir entre las frecuencias más baratas.

2.3. Análisis de la complejidad

Para realizar este análisis primero es necesario calcular cuántos intervalos puede llegar a haber en el conjunto solución (desde ahora 'CS'), este valor es, en el peor caso, $2n-1$, siendo n el número de frecuencias.

Esto se deduce de analizar las posibles entradas para el algoritmo, supongamos $n=1$. El intervalo de la frecuencia será el único elemento del CS y verifica $2n - 1 = 2 \cdot 1 - 1 = 1$.

Si ahora tomamos $n=2$, entonces dado una frecuencia (la más barata), la otra (más cara) puede que se solape o no. Si no lo hace, entonces el CS tendrá ambos intervalos incluidos. Si se solapa, 1) está incluido, con lo cual el CS será contendrá solo a la primer frecuencia; 2) arranca o termina mientras la otra está disponible, dejando al CS con únicamente el pedazo de intervalo que no se solapa y el otro entero; o bien 3) arranca y termina antes y después de que arranque y termina el primero, logrando que el CS tenga un primer intervalo de la segunda frecuencia, hasta que arranca la primera, que entra por completo y luego un segundo intervalo de la segunda frecuencia que empieza cuando termina la primera. Esto se extiende para todo n con los mismos resultados, solo que puede agregarse un 4) completar un gap.

Hasta el caso 3) se ve claramente que el máximo número de intervalos es el propuesto (agregar un intervalo a izquierda y uno a derecha). Para el caso 4) hay que tener en cuenta que el paso anterior fue un caso de no solapamiento, es decir que si intercambiara la frecuencia que completa gaps con la no solapada, entonces entraría en el caso 2) o en el caso 3), verificando que a lo sumo se llega a $2n-1$ intervalos.

El algoritmo divide tiene complejidad $T(n)=2T(n/2)+O(\text{conquer})$. Donde conquer va a recorrer, en el peor caso, dos vectores cuyas longitudes suman n , y aplicar operaciones que toman $O(1)$ para cada una de ellas. El vector que va construyendo con el CS también toma $O(n)$. Estas complejidades se suman y por propiedades de O se obtiene $O(n)$.

Reemplazando en la ecuación, $T(n)=2T(n/2)+O(n)$ que es la misma ecuación de recurrencia que MergeSort, que por Teorema Maestro se deduce que tiene complejidad $O(n \cdot \log(n))$, como pretendíamos.

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

2.5.2. Modificación del algoritmo

Debido a que nuestro algoritmo no considera como caso específico el ordenar distintas frecuencias con el mismo valor, al momento de elegir intervalos se podrían realizar cortes innecesarios. Es decir, no siempre se elige la cantidad mínima posible en estos casos.

Como consideramos que esto puede tardar un tiempo no despreciable, modificamos el operador $<$ para frecuencias, de modo que cuando tiene dos frecuencias del mismo valor ingrese primero la que tenga el comienzo antes y, en caso de comenzar en simultáneo, ingrese primero la frecuencia de mayor tamaño.

Buena justificacion! me gusta :)

Aca poner el codigo que se deberia modificar.

Aca va el grafico de esto

Comentar si paso lo que esperabamos o no.

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 ?? se pueden ver todas las casillas que están cubiertas por un sólo caballo.

3.2. Resolución propuesta y justificación

Esto va en algun lado?

Ya no se ni donde van las cosas! JAJAJA

Para la resolución de este ejercicio, se pedía un algoritmo de backtracking. La solución que presentamos tiene incluídas estrategias para que el algoritmo resuelva en un tiempo razonable los problemas medianos.

Vale aclarar que si el tablero ya estaba cubierto de antemano, simplemente había que chequearlo, tomando $O(n^2)$.

En primera instancia decidimos armar un algoritmo de backtracking "bobo", que revise todas las posibles combinaciones de tableros y nos devuelve la más óptima para este problema.

Simplemente se fija cuántos caballos necesita colocar para cubrir el tablero si no coloca un caballo en alguna posición y luego se fija cuántos harán falta si lo pone en la misma posición. Este método de "fuerza bruta" tiene una complejidad de $O(2^{n^2-k})$ siendo n la dimensión del tablero y k la cantidad de caballos preubicados. Esto es para cada posición sin caballos, ver que pasa si tomo alguna de las dos posibles decisiones.

Para disminuir la cota de complejidad, se plantearon podas y estrategias, es decir, determinar si vale la pena o no seguir revisando alguna rama del árbol de soluciones que propone esta técnica de programación.

La primera poda fue la más intuitiva, si tenemos una solución con k caballos extra agregados, y analizando otra rama llegamos a tener que necesitar agregar un caballo a una subsolución de $k-1$ caballos (o sea que tendría k caballos), entonces no nos interesa seguir revisandola, pues tenemos una solución que es igual o más óptima con k caballos.

SHAMUSHO

La segunda estrategia fue plantear si en algún momento sabíamos que debíamos agregar o no un caballo en un determinado casillero. Entonces, salteamos las k posiciones de los caballos preubicados y además salteamos aquellas posiciones que, estando atacadas, si le pusieramos un caballo, estarían atacando a casilleros que ya están siendo atacados por otros caballos.

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