



DEPARTAMENTO  
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

# Trabajo Práctico 1

Informe y análisis de resultados.

Algoritmos y Estructuras de Datos III

Integrante	LU	Correo electrónico
Benitti, Raul	592/08	raulbenitti@gmail.com
Mengarda, Lucas		l.j.mengarda@gmail.com
Scarpino, Gino	392/08	gino.scarpino@gmail.com
Vallejo, Nicolás		nico_pr08@hotmail.com

Reservado para la cátedra

Instancia	Docente	Nota
Primera entrega		
Segunda entrega		



**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>

# Índice

<b>Introducción</b>	<b>3</b>
<b>1. Problema 1</b>	<b>4</b>
1.1. Introducción . . . . .	4
1.2. Desarrollo . . . . .	4
1.2.1. Análisis de complejidad . . . . .	4
1.3. Conclusiones . . . . .	4
<b>2. Problema 2: Sensores defectuosos</b>	<b>5</b>
2.1. Introducción . . . . .	5
2.2. Desarrollo . . . . .	5
2.2.1. Correctitud . . . . .	5
2.2.2. Análisis de complejidad . . . . .	5
2.3. Resultados . . . . .	5
2.4. Conclusiones . . . . .	5
<b>3. Problema 3: La caja en el plano</b>	<b>6</b>
3.1. Introducción . . . . .	6
3.2. Desarrollo . . . . .	7
3.3. Algoritmo . . . . .	8
3.3.1. Correctitud . . . . .	9
3.3.2. Análisis de Complejidad . . . . .	11
3.4. Pruebas . . . . .	12
3.5. Conclusiones . . . . .	12
<b>4. Problema 1</b>	<b>13</b>
4.1. Introducción . . . . .	13
4.2. Desarrollo . . . . .	13
4.2.1. Análisis de complejidad . . . . .	13
4.3. Conclusiones . . . . .	13

## Introducción

En el presente trabajo práctico se intenta resolver ciertos problemas brindados por la cátedra, por medio de algoritmos que resuelvan dichos problemas. Se decidió implementar dichas soluciones usando el lenguaje *Java* y *C++*. Junto con las implementaciones, se adjunta un informe que especifica los detalles sobre el desarrollo de las soluciones, así como pruebas, cálculos de complejidad temporal, mediciones, etc.

Cabe aclarar, que dada la naturaleza de los problemas planteados, se decidió calcular y expresar la complejidad de los algoritmos involucrados, usando el **modelo de costos uniforme**.

# **1. Problema 1**

## **1.1. Introducción**

## **1.2. Desarrollo**

### **1.2.1. Análisis de complejidad**

## **1.3. Conclusiones**

## **2. Problema 2: Sensores defectuosos**

### **2.1. Introducción**

Se tiene un conjunto de sensores con sus respectivos intervalos de tiempo de medición y se conoce el número de medición que falla. Se desea conocer el id del sensor que falló.

### **2.2. Desarrollo**

Encaramos este problema, teniendo los tiempos de medición, rearmando la secuencia de mediciones de los sensores. Consiguiendo ésto, la solución al problema sería fijarse en la secuencia, el id que figura en la posición de medición que falló.

Para armar la secuencia de mediciones, decidimos ir guardando para cada sensor en que instante tendría que volver a realizar una medición.

#### **2.2.1. Correctitud**

#### **2.2.2. Análisis de complejidad**

### **2.3. Resultados**

### **2.4. Conclusiones**

### 3. Problema 3: La caja en el plano

#### 3.1. Introducción

En este caso, el problema que se nos presentó para resolver fue el siguiente:

Se tienen  $n$  puntos en el plano con coordenadas enteras y se tiene además una caja, representada por un rectángulo de dimensiones dada. La caja puede ubicarse en cualquier lugar del plano, pero no puede rotarse, es decir, la base de la caja debe quedar paralela al eje  $x$  y la altura de la caja debe quedar paralela al eje  $y$ . Un punto sobre un borde de la caja se considera dentro de la misma. Se desea ubicar la caja de manera tal que la cantidad de puntos que queden dentro de la caja sea máxima. El algoritmo que resuelva este problema debe hacerlo con una complejidad de a lo sumo  $O(n^3)$ .

Es oportuno realizar alguna aclaraciones. En primer lugar, que el problema siempre tiene solución óptima, es decir que siempre habrá una ubicación de la caja que cubra mayor o igual cantidad de puntos que todas las demás. En segundo lugar, que la solución óptima no siempre es única.

A continuación se provee un ejemplo de una posible instancia del problema con una posible solución y que además sirve para mostrar las aclaraciones hechas anteriormente:

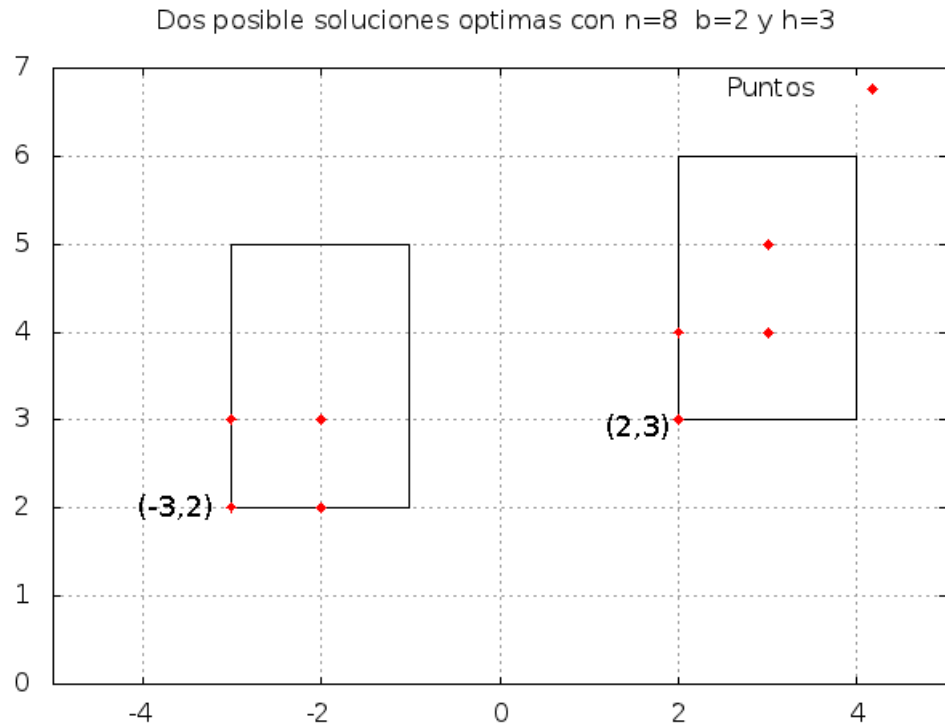


Figura 1: Aquí se observa como existe más de una solución óptima para el problema, puesto que tanto  $(-3, 2)$  como  $(2, 3)$  cubren la máxima cantidad de puntos posibles, en este caso cuatro.

Para resolver este problema hemos asumido que por lo menos se tiene algún punto en el plano (es decir que  $n > 0$ ) y que además los puntos que se proveen por entrada no están repetidos.

### 3.2. Desarrollo

Dado el problema a resolver, la primera noción que se extrajo al momento de analizarlo fue el hecho de que siempre existe una solución óptima, aunque no siempre es única.

Una primera aproximación intuitiva para la resolución consistió en ir posicionando la esquina inferior izquierda de la caja en todos los lugares posibles de un área acotada por cuatro puntos, cada uno el más extremo hacia la izquierda, la derecha, arriba o abajo. Si bien parece lógico encontrar la solución deseada de esta manera, el algoritmo sería muy complejo (y en particular la complejidad estaría condicionada por la lejanía entre los puntos extremos), por lo que resultaba necesario acotar de alguna manera las posibles posiciones donde se evaluaría la esquina inferior izquierda de la caja.

Una segunda aproximación consistió en ir probando cada punto en los cuatro vértices de la caja y evaluar cuántos de los demás puntos entrarían en ella. Sin embargo, esta idea no resultó ser fructífera: solamente estábamos evaluando potenciales soluciones que cumplieran con tener por lo menos un punto en el vértice y no es cierto que siempre existe una solución óptima que tenga por lo menos un punto en el vértice.

Luego de analizar más detenidamente el contexto del problema, se comprendió que a partir de un posicionamiento de la esquina inferior izquierda de la caja con la condición de que por lo menos un punto quedara cubierto por ella, uno podría encontrar otro posicionamiento que cumpliera que por lo menos un punto cubierto por ella se encuentre en el borde inferior de la caja y otro en el borde izquierdo de la caja (puede darse el caso que sean el mismo punto, si es así el punto estaría ubicado en el vértice inferior izquierdo de la caja) y que la cantidad de puntos dentro de la caja de este posicionamiento es mayor o igual que la cantidad del primer posicionamiento.

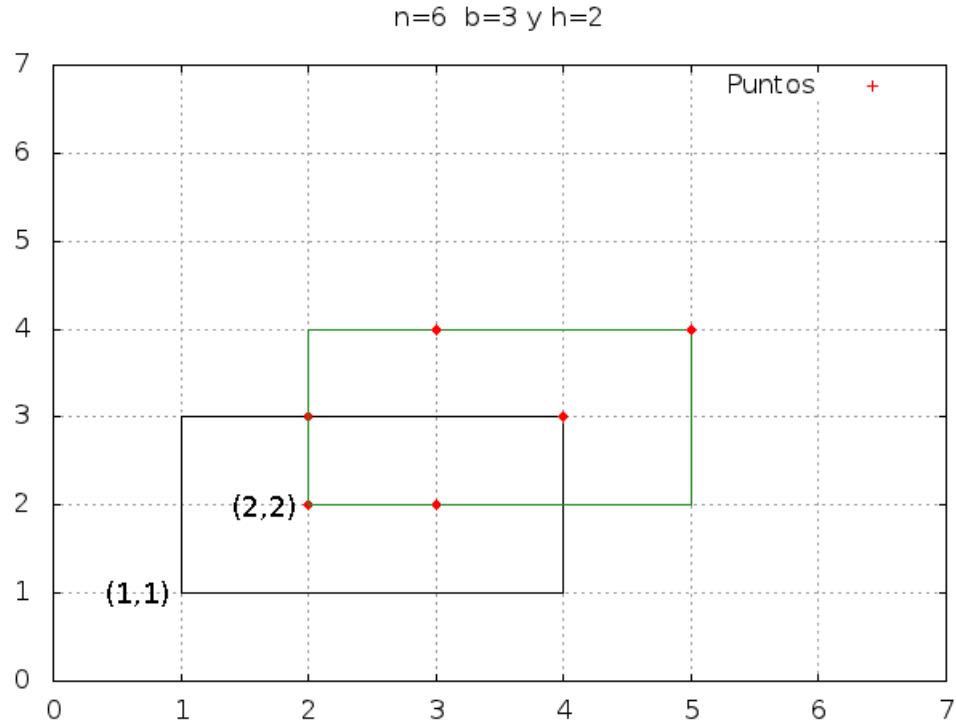


Figura 2: Aquí se observa como a partir de la posición  $(1,1)$  de una caja se puede hallar otra (la  $(2,2)$ ) que tenga por lo menos un punto en el borde, los mismos puntos que la  $(1,1)$  y que puede llegar a cubrir más puntos.

De aquí se extrae que siempre hay por lo menos una solución óptima que cumpla con estos criterios, puesto que la solución óptima debe cubrir por lo menos un punto. En particular, esa solución óptima cumple con que el componente  $x$  de su esquina inferior izquierda se corresponde con la coordenada  $x$  de alguno de los puntos que se tienen en el plano y que el componente  $y$  de su esquina inferior izquierda se corresponde con la coordenada  $y$  de alguno de los puntos que se tienen en el plano (y que no necesariamente son el mismo punto). Una demostración de todo esto se puede encontrar en la sección de Correctitud de este problema.

De esta manera hemos reducido el conjunto de posibles soluciones a uno más pequeño y que sabemos con seguridad que contiene una solución óptima.

Luego, las potenciales soluciones son las esquinas inferiores izquierdas que cumplen que su coordenada  $x$  se corresponde con la componente  $x$  de alguno de los puntos en el plano y que su coordenada  $y$  se corresponde con la componente  $y$  de alguno de los puntos en el plano.

Así, hay a lo sumo  $n^2$  candidatos a solución a analizar que se obtienen de combinar la coordenada  $x$  de cada punto con las coordenadas  $y$  de todos los puntos. En particular, nuestro algoritmo recorre estos candidatos a solución y devuelve el candidato que más puntos cubra, siendo así una solución óptima.



### 3.3. Algoritmo

Queremos hacer un par de aclaraciones:  $n$ ,  $b$  y  $h$  son de tipo *int* y representan la cantidad de puntos, la base de la caja y la altura de la caja respectivamente. A su vez, *puntosX* contiene en la posición  $i$  la coordenada  $x$  del  $i$ -ésimo punto de la entrada. Análogamente, *puntosY* contiene en la posición  $i$  la coordenada  $y$  del  $i$ -ésimo punto de la entrada. Por lo tanto, la longitud de ambos vectores es  $n$ , la cantidad de puntos en el plano.

El algoritmo recorre los candidatos a solución que cumplen que su coordenada  $x$  coincide con la de algún punto del sistema y que su coordenada  $y$  coincide con la de algún punto del sistema (que puede no ser el mismo que el de la coordenada  $x$ ).

Esto se basa en el hecho de que en dicho conjunto de puntos seguro existe por lo menos una solución óptima. Se provee la demostración de esta afirmación en la sección Correctitud de este problema.

---

#### Algoritmo 3.1

---

```

RESOLVER( $n, b, h, \text{vector} < \text{int} > \text{puntosX}, \text{vector} < \text{int} > \text{puntosY}$ )
1  int  $cant = 0$ 
2  int  $m = 0$ 
3  int  $resx = \text{puntosX}[0]$ 
4  int  $resy = \text{puntosY}[0]$ 
5  for int  $i$  desde 0 hasta  $n$ 
6      for int  $j$  desde 0 hasta  $n$ 
7           $m = \text{cuantosEntran}(n, b, h, \text{puntosX}[i], \text{puntosY}[j], \text{puntosX}, \text{puntosY})$ 
8          if ( $cant < m$ )
9               $cant = m$ 
10              $resx = \text{puntosX}[i]$ 
11              $resy = \text{puntosY}[j]$ 
12  return  $cant, resx, resy$ 

```

---



---

#### Algoritmo 3.2

---

```

CUANTOSENTRAN( $n, b, h, x, y, \text{vector} < \text{int} > \text{puntosX}, \text{vector} < \text{int} > \text{puntosY}$ )
1  int  $cant = 0$ 
2  for int  $i$  desde 0 hasta  $n$ 
3      if ( $(x \leq \text{puntosX}[i] \leq x + b) \text{ AND } (y \leq \text{puntosY}[i] \leq y + h)$ )
4           $cant++$ 
5  return  $cant$ 

```

---

**3.3.1. Correctitud**

Sean  $s = (s_x, s_y)$  una coordenada entera que representa la esquina inferior izquierda de la caja,  $P$  el conjunto de los puntos en el plano,  $n$  la cantidad de puntos en el plano,  $b$  la base de la caja y  $h$  la altura de la caja. Y sea  $P_s = \{p_1 \dots p_k\}$  el conjunto de los puntos que quedan cubiertos por la caja cuya esquina inferior izquierda está en  $s$ .

Definimos una solución óptima para este problema como aquel  $s$  que cumple que  $\#(P_s) \geq \#(P_r)$  para todo  $r$  coordenada entera.

En primer lugar, veamos que  $\exists$  un  $t$  solución óptima tal que  $t_x = e_x \wedge t_y = f_y$  para algunos  $e, f \in P$  (no es necesario que  $e = f$ ).

Para ello, veamos que siempre existe una solución óptima con por lo menos un punto en el borde de la caja.

Sea  $s$  una coordenada. Supongamos  $P_s = \{p_1 \dots p_k\} \neq \emptyset$ , es decir que la caja representada por  $s$  cubre por lo menos un punto.

Para todo  $p \in P_s$  vale que

$$\begin{aligned} s_x &\leq p_x \leq s_x + b \wedge \\ s_y &\leq p_y \leq s_y + h \end{aligned}$$

puesto que están cubiertos por la caja.

Ahora bien, sean  $a = \min_{i:1\dots k}((p_i)_x)$  y  $c = \min_{i:1\dots k}((p_i)_y)$ .

Definamos ahora  $s'$  tal que  $s'_x = a$   $s'_y = c$ .

Entonces, para todo  $i$  desde 1 hasta  $k$ :

$$\begin{aligned} s'_x = a &= \min_{i:1\dots k}((p_i)_x) \leq (p_i)_x \wedge \\ s'_y = c &= \min_{i:1\dots k}((p_i)_y) \leq (p_i)_y \end{aligned}$$

Es decir,

$$\begin{aligned} s'_x &\leq p_x \wedge \\ s'_y &\leq p_y \end{aligned}$$

para todo  $p \in P_s$ .

Además, por como definimos  $s'$ :

$$\begin{aligned} s_x &\leq s'_x \wedge \\ s_y &\leq s'_y \end{aligned}$$

Por lo tanto, para todo  $p \in P_s$ :

$$\begin{aligned} s_x \leq s'_x \leq p_x \leq s_x + b \leq s'_x + b \wedge \\ s_y \leq s'_y \leq p_y \leq s_y + h \leq s'_y + h \end{aligned}$$

Es decir que:

$$\begin{aligned} s'_x \leq p_x \leq s'_x + b \wedge \\ s'_y \leq p_y \leq s'_y + h \end{aligned}$$

y eso quiero decir que para todo punto  $p \in P_s$  vale que  $p \in P_{s'}$ .

Luego,  $P_s \subseteq P_{s'}$ .

Es decir que  $\#P_s \leq \#P_{s'}$ .

Notar que no vale la vuelta, puesto que al reubicar la caja se pueden estar cubriendo puntos que antes no estaban cubiertos.

En particular, si  $s$  es una solución óptima, vale que  $\#(P_s) \geq \#(P_r)$  para todo  $r$  coordenada entera (recordamos que en la plano hay por lo menos un punto y que necesariamente la solución óptima tiene que cubrir por lo menos un punto).

Luego,  $\#P_s \geq \#P_{s'}$ . Pero como  $\#P_s \leq \#P_{s'}$ , entonces:

$$\#P_s = \#P_{s'}$$

y entonces  $s'$  también es solución óptima. Como siempre se puede encontrar una solución óptima  $s$ , entonces también siempre se puede encontrar una solución óptima  $s'$ . Es decir que hemos encontrado una solución óptima que cumple que tiene por lo menos un punto en el borde.

Definamos ahora el conjunto de coordenadas

$$Z = \{z | z_x = e_x \text{ para algún } e \in P \wedge z_y = f_y \text{ para algún } f \in P\}$$

Es decir que  $Z$  es el conjunto de coordenadas que cumplen que su componente  $x$  se corresponde con el de algún punto en el plano y que su componente  $y$  se corresponde con la algún punto en el plano (no necesariamente el mismo punto).

Es fácil observar que de la manera en que lo definimos,  $s' \in Z$ .

Esto es porque  $s'_x = \min_{p \in P_{s'}}(p_x) \wedge s'_y = \min_{p \in P_{s'}}(p_y)$ .

Además como siempre puedo encontrar un  $s'$  con estas condiciones que sea solución óptima, eso quiere decir que seguro existe una solución óptima en  $Z$ .

Por lo tanto,  $\exists$  una coordenada  $t$  solución óptima tal que  $t_x = e_x \wedge t_y = f_y$  para algunos  $e, f \in P$  (no es necesario que  $e = f$ ), que es lo que queríamos demostrar.

Ahora veamos como nuestro algoritmo encuentra una solución correcta.

Es fácil ver que en efecto el conjunto  $Z$  se puede formar combinando las componentes  $x$  de cada punto en el plano con las componentes  $y$  de los todos puntos en el plano. Como ya hemos demostrado, en ese conjunto seguro hay alguna solución óptima  $s$  que en particular cumple que para todo  $z \in Z$   $\#P_s \geq \#P_z$ .

El algoritmo recorre entonces todos los elementos  $z$  del conjunto  $Z$ . Mediante la función *cuantosEntran* se cuenta la cantidad de puntos que entran en la caja que tiene como esquina inferior izquierda a  $z$ , es decir que devuelve  $\#P_z$  y lo conserva como solución parcial si y sólo si es la que mayor puntos cubre hasta ese momento. Es decir que en cada iteración la solución parcial (llamémosla  $v$ ) cumple que  $\#P_v$  es mayor que la todos los puntos  $z$  que se evaluaron hasta ese momento.

Luego, cuando se iteraron todos los elementos de  $Z$  la solución  $s$  que tenemos cumple que  $\#P_s \geq \#P_z$  para todo  $z \in Z$ .

Pero como sabemos que en ese conjunto existe una solución óptima, entonces  $s$  también lo es. En particular, lo que puede ocurrir es que devuelva esa solución óptima que demostramos que existe en el conjunto  $Z$  (es decir la que tiene puntos en los bordes de la caja) o que devuelva otra que no cumpla con esa condición, pero que al cumplir que  $\#P_s \geq \#P_z$  también es solución óptima.

□

### 3.3.2. Análisis de Complejidad

Analicemos en primer lugar la complejidad de *cuantosEntran*. Asignarle un valor a la variable *int cant* cuesta tiempo constante, es decir  $\mathbf{O}(1)$ . A su vez, evaluar la condición del If también cuesta tiempo constante. Esto es porque realizar la comparación de *int's* es constante y, de acuerdo a la documentación de C++, obtener un elemento de un vector mediante el operador `[]` también lo es. Por lo tanto, evaluar la condición del If es  $\mathbf{O}(1)$ . Luego, se entre o no al cuerpo del If, en cada iteración del ciclo se ejecuta algo de costo constante. Como el ciclo se itera  $n$  veces (siendo  $n$  la cantidad de puntos que hay en el plano), el algoritmo *cuantosEntran* tiene un costo de

$$\mathbf{O}(1 + n*1)$$

con lo cual tiene una complejidad temporal  $\mathbf{O}(n)$ , es decir lineal en la cantidad de puntos.

Analicemos ahora *resolver*. Las primeras cuatro asignaciones se realizan en tiempo constante (incluso las que obtienen un valor del vector con el operador `[]`), es decir en  $\mathbf{O}(1)$ .

Veamos que ocurre en el cuerpo del segundo For. Se ejecuta *cuantosEntran*, que ya sabemos que lo hace en tiempo lineal, y lo que devuelve se lo asigna a una variable (en tiempo constante). Luego se hace una comparación en  $\mathbf{O}(1)$ .

Ahora bien, si no entra al cuerpo del If no se hace nada más y se pasa a la próxima iteración. En este caso el costo de la iteración fue de  $\mathbf{O}(n+2)$ , con lo cual tiene

una complejidad temporal de  $\mathbf{O}(n)$ .

En cambio, si entra al cuerpo del If, se realizan tres asignaciones, todas en tiempo constante. Luego lo que está adentro del cuerpo del If tiene un costo de  $\mathbf{O}(1)$ . Es decir que en este caso también una iteración cuesta  $\mathbf{O}(n)$ .

De aquí se extrae que lo que ocurre en cada iteración de este For tiene un costo lineal. Como este For se ejecuta  $n$  veces, su complejidad temporal será  $\mathbf{O}(n^2)$ .

Entonces, el primer For ejecuta en cada iteración algo con un costo cuadrático. Como se itera  $n$  veces se obtiene que el costo de este For es  $\mathbf{O}(n^3)$ . Es decir que el algoritmo cuesta

$$\mathbf{O}(1+1+1+ n^3)$$

con lo cual tiene un complejidad temporal de  $\mathbf{O}(n^3)$ , es decir cúbica en la cantidad de puntos, cumpliendo entonces con la cota que se nos exigía en el enunciado del problema.

### 3.4. Pruebas

### 3.5. Conclusiones

## **4. Problema 1**

### **4.1. Introducción**

### **4.2. Desarrollo**

#### **4.2.1. Análisis de complejidad**

### **4.3. Conclusiones**