



## DEPARTAMENTO DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

# Trabajo Práctico III

Algoritmos y Estructura de Datos III  
Primer Cuatrimestre de 2015

Integrante	LU	Correo electrónico
Iván Arcuschin	678/13	iarcuschin@gmail.com
Martín Jedwabny	885/13	martiniedva@gmail.com
José Massigoge	954/12	jmmassigoge@gmail.com
Lucas Puterman	830/13	lucasputerman@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>

# Índice

<b>1. Introducción</b>	<b>3</b>
<b>2. Ejercicio 1 - Demostración</b>	<b>4</b>
2.1. Definiciones . . . . .	4
2.2. Ejercicio A . . . . .	4
2.3. Ejercicio B . . . . .	6
2.4. Ejercicio C . . . . .	6
<b>3. Ejercicio 2 - Algoritmo exacto</b>	<b>7</b>
3.1. Ejercicio A . . . . .	7
3.1.1. Estrategia . . . . .	7
3.1.2. Podas . . . . .	7
3.1.3. Pseudocódigo . . . . .	8
3.2. Ejercicio B . . . . .	8
3.3. Ejercicio C . . . . .	10
<b>4. Ejercicio 3 - Heurística constructiva golosa</b>	<b>13</b>
4.1. Ejercicio A . . . . .	13
4.2. Pseudocódigo . . . . .	13
4.2.1. Implementación sobre listas de adyacencia . . . . .	13
4.2.2. Implementación sobre matriz de adyacencia . . . . .	14
4.3. Ejercicio B . . . . .	14
4.4. Ejercicio C . . . . .	15
4.5. Ejercicio D . . . . .	17
4.5.1. Experimentación sobre grafo completo . . . . .	17
4.5.2. Experimentación sobre el complemento del grafo completo . . . . .	23
<b>5. Ejercicio 4 - Heurística de búsqueda local</b>	<b>30</b>
5.1. Ejercicio A . . . . .	30
5.1.1. Algoritmo implementado . . . . .	30
5.1.2. Procedimiento BFS modificado: . . . . .	30
5.1.3. Primer Criterio de Vecindad . . . . .	31
5.1.4. Segundo Criterio de Vecindad . . . . .	31
5.2. Ejercicio B . . . . .	33
5.3. Ejercicio C . . . . .	34
5.3.1. Tiempos de ejecución . . . . .	35
5.3.2. Calidad de la solución . . . . .	36
<b>6. Ejercicio 5 - Metaheurística GRASP</b>	<b>40</b>
6.1. Ejercicio A . . . . .	40
6.1.1. Idea general . . . . .	40
6.1.2. Criterios de parada . . . . .	40
6.1.3. Selección de lista de candidatos (RCL) . . . . .	40
6.1.4. Pseudocódigo . . . . .	41
6.2. Ejercicio B . . . . .	42
6.2.1. Tiempos de ejecución . . . . .	43
6.2.2. Calidad de las respuestas . . . . .	46
<b>7. Ejercicio 6 - Experimentación final</b>	<b>48</b>
7.1. Experimentación con Grafos completos . . . . .	48
7.1.1. Tiempos de ejecución . . . . .	49
7.1.2. Calidad de las soluciones . . . . .	50
7.2. Experimentación con Complemento de Grafos completos . . . . .	50

## Sección ÍNDICE

---

7.2.1. Tiempos de ejecución . . . . .	50
7.2.2. Calidad de las soluciones . . . . .	51
7.3. Experimentación con Caminos . . . . .	51
7.3.1. Tiempos de ejecución . . . . .	51
7.3.2. Calidad de las soluciones . . . . .	53
7.4. Experimentación con Ciclos . . . . .	53
7.4.1. Tiempos de ejecución . . . . .	53
7.4.2. Calidad de las soluciones . . . . .	54
<b>8. Conclusión</b>	<b>56</b>
<b>A. Implementación algoritmo exacto</b>	<b>57</b>
<b>B. Implementación algoritmo goloso</b>	<b>60</b>
B.1. Implementación sobre matriz de adyacencia . . . . .	60
B.2. Implementación sobre listas de adyacencia . . . . .	61
<b>C. Implementación algoritmo de búsqueda local</b>	<b>64</b>
<b>D. Implementación metaheurística GRASP</b>	<b>70</b>

## 1. Introducción

Este Trabajo Práctico se centra alrededor del problema de *Conjunto Independiente Dominante Mínimo* (CIDM) que consiste en hallar un conjunto independiente dominante de  $G$  con mínima cardinalidad.

Como se sabe<sup>1</sup>, este problema pertenece a la categoría de problemas **NP**. Esto quiere decir, entre otras cosas, que no se conoce hasta el día de hoy una solución exacta en tiempo polinomial que lo resuelva.

Es en el contexto de este tipo de problemas que cobra sentido el análisis de los algoritmos de solución aproximada. Dichos algoritmos no garantizan encontrar la mejor solución (de hecho, en algunos casos pueden dar soluciones bastante malas), pero a cambio podemos obtener dicha solución en tiempo polinomial.

El objetivo de este Trabajo Práctico entonces, es estudiar la factibilidad de distintos algoritmos Heurísticos para el problema CIDM. En particular nos interesa implementar y experimentar con los algoritmos **Constructivo Goloso**, de **Búsqueda Local** y **GRASP**. Este último pertenece a la categoría de algoritmos de Metaheurística ya que combina los dos primeros (Goloso y Búsqueda Local) en un algoritmo.

Además, para tener tiempos de ejecución y soluciones de referencia se implementó un algoritmo que resuelve el problema de forma exacta, pero tomando tiempo exponencial.

Por último, se realizaron distintas experimentaciones para comparar los tiempos de ejecución y la calidad de las soluciones de los algoritmos mencionados.

---

<sup>1</sup>Referencia: [https://en.wikipedia.org/wiki/Dominating\\_set](https://en.wikipedia.org/wiki/Dominating_set)

## 2. Ejercicio 1 - Demostración

### 2.1. Definiciones

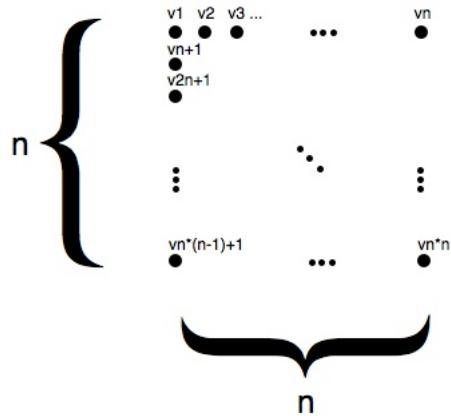
Sea  $G = (V, E)$  un grafo simple:

- Definición 1:  $D \subseteq V$  es un **conjunto dominante** (CD)  $\iff \forall v \in V, v \in D \text{ ó } \exists w \in D \text{ tal que } (v, w) \in E$  (tiene un vecino en D).
- Definición 2:  $D \subseteq V$  es un **conjunto independiente** (CI)  $\iff \forall v, w \in D, (v, w) \notin E$ .
- Definición 3:  $D \subseteq V$  es un **conjunto independiente dominante** (CID)  $\iff D$  es dominante e independiente.
- Definición 4:  $D \subseteq V$  es un **conjunto independiente dominante mínimo** (CIDM)  $\iff D$  es el conjunto independiente dominante de  $V$  con menos nodos. Es decir que  $\forall D' \subseteq V$  tal que  $D'$  es independiente dominante,  $\#(D) \leq \#(D')$ .
- Definición 5:  $D \subseteq V$  es un **conjunto independiente maximal** (CIMax)  $\iff D$  es independiente y  $\nexists D' \subseteq V$  independiente tal que  $D \subset D'$  ( $\subset$  estricto).

### 2.2. Ejercicio A

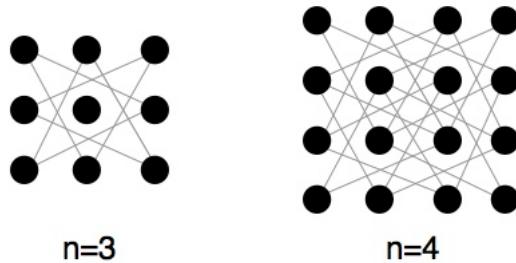
*Relacionar el problema de CIDM con el problema 3 del TP 1 (i.e., similitudes y diferencias):*

Si definimos a  $G = (V, E)$  grafo simple para que represente al tablero de ajedrez que se forma en el ejercicio 3 del tp1, los nodos serían una cuadricula de la pinta (sin considerar los ejes, eso lo hacemos después):



Donde 'n' es el parámetro del ejercicio que indica el tamaño de lado del tablero.

Ahora, como el ejercicio consiste en poner la mínima cantidad de caballos para que todas las casillas tengan un caballo o bien estén amenazadas, podemos representar las 'amenazas' como los ejes del grafo, por ejemplo:



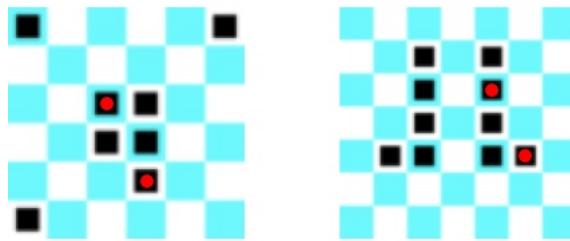
*Observación: como cualquier posición x que amenaza a otra posición y sería amenazada si pongo un caballo en y, podemos representar el problema con un grafo simple y no un digrafo.*

Entonces los ejercicios son similares ya que lo que estamos buscando en el ejercicio del TP1 es la mínima cantidad de caballos para que todas las posiciones (nodos) tengan un caballo o estén amenazadas (dominadas).

Es decir, si los caballos son un conjunto D de vértices en V, queremos hallar  $\#(D)$  tal que:

- $\forall v \in V, v \in D \text{ o } \exists w \in D \text{ tal que } (v,w) \in E \iff D \text{ es dominante.}$
- $\forall D' \subseteq V \text{ tal que } D' \text{ es dominante, } \#(D) \leq \#(D') \iff D \text{ es mínimo entre los conjuntos dominantes de } V.$

Por lo tanto el problema consiste en hallar un conjunto dominante mínimo pero, a diferencia del problema en este TP, puede o no ser independiente. Esto se debe a que existen respuestas para el problema del TP1 donde hay caballos en ciertas posiciones que se amenazan entre sí:



*Observación: estos graficos fueron extraidos de la página: <http://home.earthlink.net/~morgenstern/solution/knsols1.htm>. El primero representa una solución óptima del ejercicio del TP1 con un tablero de 6x6 y la segunda imagen es para un tablero de 7x7. En ambas, se marcan con un punto negro los casilleros que tienen un caballo y, como vemos señalado en rojo, hay caballos en posiciones que se amenazan entre sí.*

Mas alla de eso, ambos problemas son de optimización combinatoria. Los conjuntos factibles son los vértices de un grafo que cumplen ciertas condiciones (si bien como dijimos antes, no son las mismas para los dos problemas) y la función de optimalidad consiste en minimizar la cantidad de vértices del conjunto.

### 2.3. Ejercicio B

*Demostrar que todo conjunto independiente maximal es un conjunto dominante:*

Sea  $G = (V, E)$  un grafo simple y  $D \subseteq V$  un conjunto independiente maximal, quiero ver que  $D$  es un conjunto dominante.

Supongamos (por absurdo) que  $D$  no es un conjunto dominante:

$$\Rightarrow \exists v \in V \text{ tal que } v \notin D \wedge \forall w \in D, (v, w) \notin E \text{ (no es vecino de ning\'un nodo en } D).$$

$$\Rightarrow D \cup \{v\} \text{ es independiente ya que como } D \text{ es independiente sucede que } \forall w \in D \cup \{v\} \ \exists x \in D \text{ tal que } (x, w) \in E.$$

$$\Rightarrow \text{Absurdo! pues } D \text{ era maximal y } \exists D' = D \cup \{v\} \text{ tal que } D' \text{ es independiente y } D \subset D', \text{ asi contradiciendo el hecho de que } D \text{ es un conjunto independiente maximal.}$$

Este absurdo vino de suponer que  $D$  no era dominante.

Por lo tanto, si  $D$  es un conjunto independiente maximal  $\Rightarrow D$  es dominante.

### 2.4. Ejercicio C

*Describir situaciones de la vida real que puedan modelarse utilizando CIDM:*

- **El turista:** tenemos un conjunto de ciudades que forman un grafo simple conexo donde los nodos son las ciudades y dos nodos estan conectados si y solo si las ciudades son vecinas. Suponiendo que queremos conocer tantas culturas distintas como sea posible, decidimos que no queremos visitar ningun par de ciudades vecinas ya que sus culturas son muy similares. Como el problema que proponemos consiste en hallar un conjunto m\'ınimo de ciudades (nodos en el mapa) tal que toda ciudad sea visitada o bien una ciudad vecina sea visitada pero no ambas (el conjunto de nodos sea independiente y dominante), podemos decir que estamos buscando un CIDM.
- **Spamear una red social:** supongamos que tenemos un virus que se ocupa de spamear Facebook de manera que un mensaje sea propagado por toda la red. Asumimos que los usuarios de Facebook son nodos en un grafo simple conexo y que dos nodos estan conectados si y solo si esos dos usuarios son amigos en la red. Ahora, el virus no quiere ser descubierto, asi que lo que debe hacer es infectar la menor cantidad de cuentas que no sean amigas, ya que si spamea demasiadas cuentas o un par de cuentas que son amigas, ser\'ia mucho m\'as f\'acil de descubrirlo. Luego, estamos buscando el m\'ınimo conjunto de nodos en la red de usuarios tal que todos los usuarios vean el mensaje (dominaci\'on) y dos amigos no sean infectados a la vez (independencia), el cual es un CIDM.
- **C\'amaras de seguridad en un barrio:** tenemos casas en un barrio que representan los nodos de un grafo simple conexo, donde dos nodos estan conectados si y solo si sus casas respectivas con vecinas. Queremos poner c\'amaras de seguridad en el barrio de manera que toda casa tenga una c\'amara de seguridad o la tenga una casa vecina, ya que el rango de cobertura de la c\'amara es amplio y puede filmar la casa donde est\'a y tambi\'en las vecinas (dominaci\'on). No queremos poner c\'amaras en dos casas vecinas (independencia) para que no sea tan notorio y arruine el paisaje. A la vez, tenemos un presupuesto limitado, por lo cual queremos poner tan pocas c\'amaras como sea posible (minimalidad). Por lo tanto, el conjunto de casas que estamos buscando para ponerles c\'amaras es un CIDM.

### 3. Ejercicio 2 - Algoritmo exacto

Diseñar e implementar un algoritmo exacto para CIDM.

#### 3.1. Ejercicio A

Explicar detalladamente el algoritmo implementado. Elaborar podas y estrategias que permitan mejorar los tiempos de resolución.

##### 3.1.1. Estrategia

El algoritmo implementado se basa en la Técnica Algorítmica de *Backtracking*.

En primer lugar notese que un conjunto dominante 'D' trivial es el conjunto de todos los nodos del grafo. Si bien este conjunto en principio no es independiente (a menos que  $X(G) = \emptyset$ ), es el conjunto dominante más grande posible (no hay ningún nodo que quede afuera, es decir:  $V(G) \setminus D = \emptyset$ ).

Luego, la idea principal del algoritmo es empezar con este primer conjunto dominante e ir sacando nodos recursivamente mientras chequeamos dominancia e independencia. Cada vez que encontramos un conjunto dominante e independiente, nos fijamos su cardinal. Si el cardinal de este nuevo conjunto es menor que el mínimo hasta ese momento nos quedamos con el nuevo y lo guardamos como mínimo. En caso de que el cardinal sea mayor, seguimos quitando nodos para encontrar un conjunto más chico.

##### 3.1.2. Podas

Antes de mostrar cuales son las podas que utilizamos en el algoritmo, veamos algunos Lemas.

**Lema 3.1.** Sea  $C$  un conjunto dominante e independiente de un grafo  $G$ . Entonces, cualquier subconjunto de  $C$  (distinto de  $C$ ) es no-dominante respecto a  $G$ .

*Demuestra*ción. Veamos que vale por absurdo. Supongamos  $H$  un subconjunto propio de  $C$  tal que  $H \subset C$  y  $H$  es dominante de  $G$ . Como  $H$  está estrictamente incluido en  $C$ , entonces  $\exists v \in C$  tal que  $v \notin H$ .

Ahora, como  $C$  es independiente, no hay nodos de  $C$  que sean adyacentes en  $G$ , en particular:  $\forall w \in C, (v, w) \notin X(G)$ .

Entonces,  $v \in V(G)$  pero  $v \notin H$  y  $\forall w \in H \subset C, (v, w) \notin X(G)$ . Luego,  $H$  no es dominante, ya que el nodo  $v$  no está en  $H$  ni tiene algún vecino que esté en  $H$ . Absurdo.  $\square$

**Lema 3.2.** Sea  $C$  un conjunto no-dominante respecto a  $G$ . Entonces, cualquier subconjunto de  $C$  es no-dominante respecto a  $G$ .

*Demuestra*ción. Trivial. Sea  $H$  un subconjunto de  $C$  ( $H \subseteq C$ ).

Como  $C$  es no-dominante,  $\exists v \in V(G)$  tal que  $v \notin C$  y  $\forall w \in C, (v, w) \notin X(G)$ .

Pero como  $H \subseteq C, v \notin H$  y  $\forall w \in H, (v, w) \notin X(G)$ .

Luego  $H$  es no-dominante respecto a  $G$ .  $\square$

Usando los resultados de estos dos lemas podemos optimizar el algoritmo cortando ramas en el árbol de recursión.

- **Poda A:** por el primer Lema, cada vez que encontramos un conjunto dominante e independiente podemos ahí mismo devolver el que tenga menor cardinal entre ese conjunto y el mínimo encontrado hasta ese momento. Esto es así ya que sabemos que cualquier subconjunto del nuevo conjunto es no-dominante.
- **Poda B:** por el segundo Lema, al momento de sacar un nodo de un conjunto podemos chequear si el subconjunto es dominante o no. En caso de que no lo sea, ni siquiera lo procesamos, ya que no es dominante y ninguno de sus subconjuntos lo será.

### 3.1.3. Pseudocódigo

```

funcion resolver:
    Creamos un vector de n elementos, llenandolo con los nodos desde 0 a n-1.
    llamamos a resolver_aux pasandole como parametro la matriz de adyacencia, el vector
        recién creado y la cantidad de nodos en el grafo.

funcion resolver_aux:
    Llamemos dom al conjunto dominante pasado por parametro.
    Llamemos cidm al conjunto dominante e independiente con menor cardinal encontrado
        hasta ahora.
    Si el cidm tiene tamaño 1 hacer
        Devuelvo cidm
    Sino hacer
        // Chequeamos si dom es independiente:
        Para i desde 0 hasta |dom| hacer:
            Para j desde i+1 hasta |dom| hacer:
                Si matriz_adyacencia[dom[i]][dom[j]] == TRUE hacer
                    dom NO es independiente
                Si es independiente hacer
                    Devolvemos el conjunto con menor cardinal entre cidm y dom.
        Para i desde 0 hasta |dom| hacer
            Crear un vector nuevo llamado copia con los mismos nodos que dom.
            Borrar el nodo en la posicion i del vector copia.
            // Chequeamos si el conjunto copia es dominante:
            Para i desde 0 hasta n hacer:
                Si i no pertenece a copia hacer
                    Bool nodo_valido = FALSE
                    Para j desde 0 hasta |copia| hacer:
                        Si copia[j] == i hacer
                            nodo_valido = TRUE
                        Si matriz_adyacencia[i][copia[j]] == TRUE hacer
                            nodo_valido = TRUE
                    Si nodo_valido == FALSE hacer
                        copia NO es dominante
                Si copia es dominante hacer
                    Hacer un llamado recursivo a resolver_aux, pasando como parametro la matriz
                        de adyacencia, cidm, copia y la cantidad de nodos en el grafo.
                    Llamemos nuevo_cidm al conjunto que devuelve el llamado recursivo.
                    Devolvemos el conjunto con menor cardinal entre cidm y nuevo_cidm.

```

## 3.2. Ejercicio B

*Calcular el orden de complejidad temporal de peor caso del algoritmo.*

Veamos el costo de las distintas partes del algoritmo.

En la función *resolver* tenemos un ciclo de  $\Theta(n)$  más un llamado al constructor de la estructura vector para construir el primer conjunto dominante que se realiza solo una vez. Este constructor tiene costo  $\Theta(n)^2$ , y dentro del ciclo solo tenemos operaciones constantes (asignación), por lo que el costo, sin contar el llamado a la función *resolver\_aux*, es de  $\Theta(n)$ .

En la función *resolver\_aux* tenemos 5 ciclos distintos, 2 anidados entre sí y otros 3 anidados entre sí. Veamos los primeros:

<sup>2</sup>Referencia: <http://www.cplusplus.com/reference/vector/vector/vector/>

```

Para i desde 0 hasta |dom| hacer:
    Para j desde i+1 hasta |dom| hacer:
        Si matriz_adyacencia[dom[i]][dom[j]] == TRUE hacer
            dom NO es independiente

```

Estos ciclos anidados corresponden al chequeo de independencia. Dentro de ellos solo hay operaciones de costo constante (una asignación y dos comparaciones) y se repiten cada uno  $|dom|$  veces, donde  $dom$  es el conjunto dominante actual. Luego, el chequeo de independencia toma:  $\mathcal{O}(|dom| * |dom|)$ , que en el peor caso es  $\mathcal{O}(n^2)$ .

Y los ciclos restantes:

```

Para i desde 0 hasta |dom| hacer
    Crear un vector nuevo llamado copia con los mismos nodos que dom.
    Borrar el nodo en la posición i del vector copia.
    // Chequeamos si el conjunto copia es dominante:
    Para i desde 0 hasta n hacer:
        Si i no pertenece a copia hacer
            Bool nodo_valido = FALSE
            Para j desde 0 hasta |copia| hacer:
                Si copia[j] == i hacer
                    nodo_valido = TRUE
                Si matriz_adyacencia[i][copia[j]] == TRUE hacer
                    nodo_valido = TRUE
                Si nodo_valido == FALSE hacer
                    copia NO es dominante
        Si copia es dominante hacer
            Hacer un llamado recursivo a resolver_aux, pasando como parámetro la matriz
            de adyacencia, cidm, copia y la cantidad de nodos en el grafo.
            Llameemos nuevo_cidm al conjunto que devuelve el llamado recursivo.
            Devolvemos el conjunto con menor cardinal entre cidm y nuevo_cidm.

```

De estos 3 ciclos anidados, el primero corresponde a la iteración sobre el conjunto  $dom$  mientras vamos probando quitar distintos nodos, y los otros 2 corresponden al chequeo de dominancia de los nuevos conjuntos generados. El primer ciclo se ejecuta  $|dom|$  veces, mientras que el segundo y el tercero  $n$  y  $|copia|$  veces respectivamente, donde  $copia$  es el nuevo conjunto generado (como se muestra en el pseudocódigo). Además dentro del primer ciclo utilizamos la función `erase` de vectores que permite borrar un elemento dado su índice y el constructor de la estructura vector por copia. Ambas operaciones tiene costo en el peor caso lineal en la cantidad de elementos del vector original<sup>3</sup>. El resto de las operaciones en los 3 ciclos es de costo constante (comparaciones y asignaciones).

Luego, la complejidad de los 3 ciclos nos queda:  $\mathcal{O}(|dom| * (|dom| + |copia| + |n| * (|copia|)))$ , que asintóticamente equivale a  $\mathcal{O}(n^3)$

Fuera de los ciclos en la función `resolver_aux` solo tenemos operaciones constantes (comparaciones y asignaciones, más la función `size` de vectores, que tiene costo constante<sup>4</sup>). Entonces, el costo de ejecutar una vez la función `resolver_aux`, sin tener en cuenta el llamado recursivo a si misma, es de  $\mathcal{O}(n^3)$ .

Ahora, como vimos en el pseudocódigo, la función `resolver_aux` parte del conjunto  $dom$  y hace en el peor caso  $|dom|$  llamados recursivos a si misma, cada uno quitando un nodo posible de  $|dom|$ .

Luego, el peor caso posible es que revisemos todas las posibles combinaciones de nodos. Es decir, pasar por `resolver_aux` la cantidad de veces del cardinal del Conjunto de Partes de  $V(G)$ . O sea, en el peor caso `resolver_aux` toma:  $\mathcal{O}(2^n * n^3)$ .

Entonces, la complejidad final del algoritmo es:

$$\mathcal{O}(2^n * n^3)$$

<sup>3</sup>Referencias: <http://www.cplusplus.com/reference/vector/vector/erase/> y <http://www.cplusplus.com/reference/vector/vector/>

<sup>4</sup>Referencia: <http://www.cplusplus.com/reference/vector/vector/size/>

### 3.3. Ejercicio C

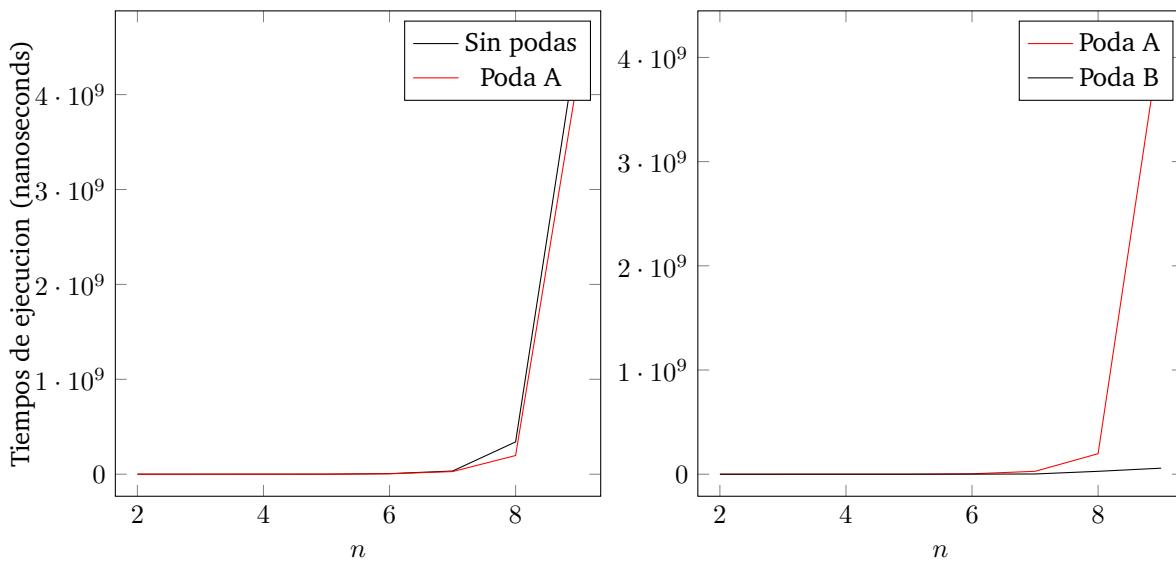
Realizar una experimentación que permita observar los tiempos de ejecución del algoritmo en función de los parámetros de la entrada y de las podas y/o estrategias implementadas.

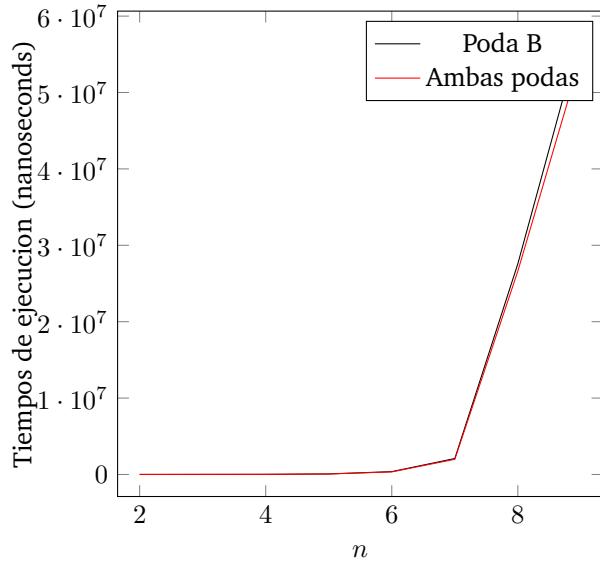
Consideraciones:

- Los casos de tests para esta experimentación corresponden a grafos generados aleatoriamente, variando la cantidad de nodos y ejes.
- Los tiempos de ejecución se midieron con la biblioteca chrono y estos fueron convertidos a nanosegundos.
- Los valores aleatorios que fuimos tomando fueron generados con una distribución uniforme para el rango requerido.
- Los detalles de los parametros son:
  - $2 \leq n \leq 9$
  - $0 \leq m \leq n * (n - 1)/2$
  - Luego, para cada eje, se eligieron 2 nodos tales que  $1 \leq v \leq n$ ,  $1 \leq w \leq n$  y  $(v, w) \notin X(G)$  donde  $G$  es el grafo que estamos armando. Esto permite que el grafo tenga exactamente  $m$  ejes (y no una cantidad menor que se podría dar al poner ejes que ya estaban).
- Para cada caso de test, se midió la ejecución del algoritmo con y sin podas.
- Para cada  $n$  se promediaron los resultados obtenidos para cada variaciones del algoritmo, con y sin podas.

Luego, queremos ver que el costo del algoritmo es, en el peor caso, el detallado en la sección Complejidad. La cota teórica calculada es  $\mathcal{O}(2^n * n^3)$ .

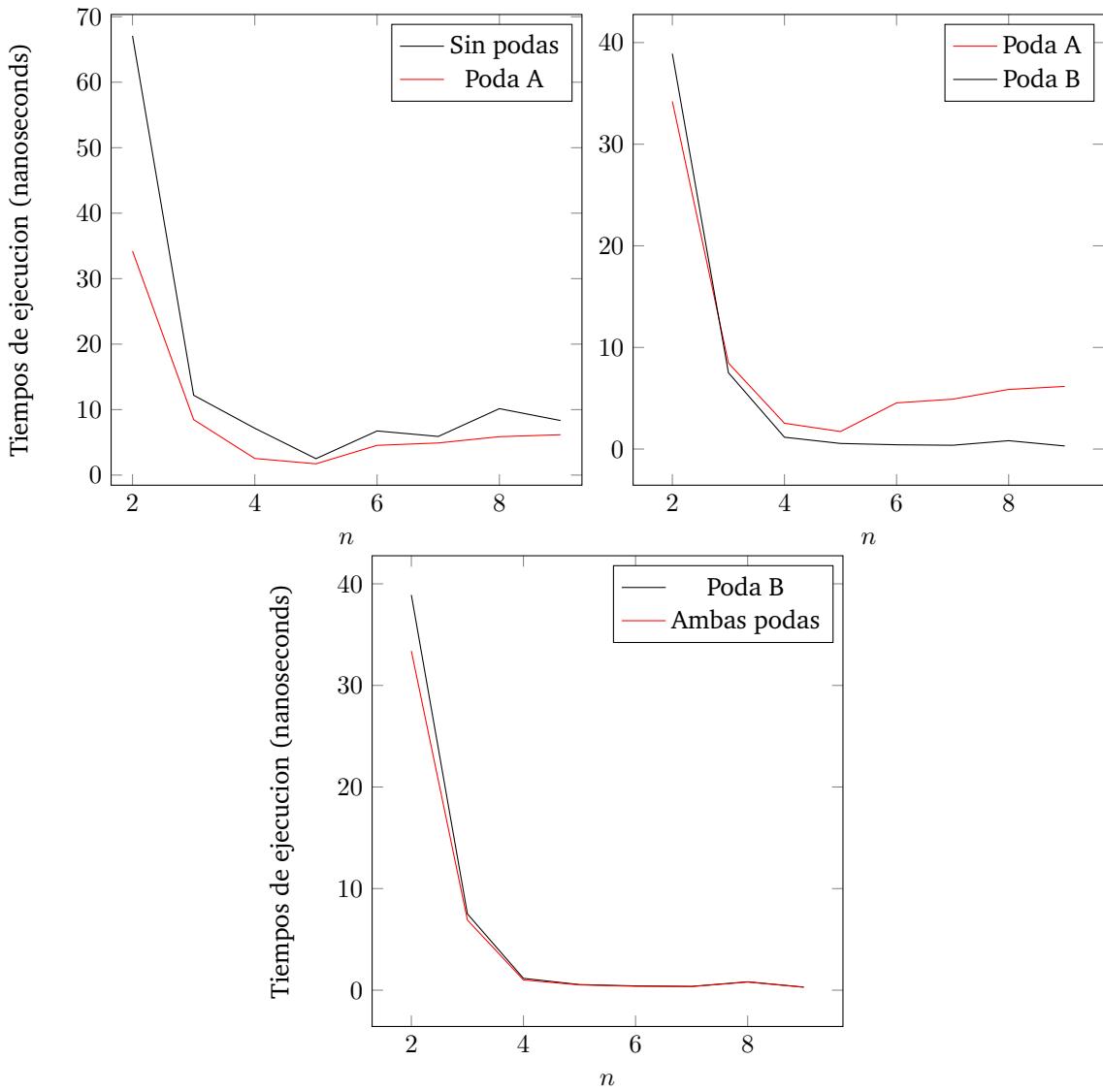
Los resultados obtenidos fueron los siguientes:





Como se puede ver en los gráficos, los tiempos medidos tienen pinta exponencial, que se condice con la complejidad teórica calculada. Además, es claro que la poda B produce tiempos mucho mejores que la poda A, lo cual en principio no se esperaba.

Veamos que pasa cuando los dividimos por  $2^n * n^3$ :



Al realizar la división, podemos ver que los gráficos tienden a una constante arriba de cero, lo cual es lo esperado. Por lo tanto, el algoritmo tendría una complejidad  $\mathcal{O}(c * 2^n * n^3)$ , donde  $c$  es la constante a la cual converge el gráfico.

## 4. Ejercicio 3 - Heurística constructiva golosa

### 4.1. Ejercicio A

Realizamos una heurística golosa para resolver el problema. Lo que hace el algoritmo es lo siguiente. Primero toma los datos de entrada del problema, y se arma la matriz o las listas de adyacencia del grafo (según la implementación elegida). Además, se crea un vector de nodos de tamaño  $n$  donde cada nodo contiene guardado su número de nodo y el grado que tiene en el grafo.

Una vez que se procesaron los datos de entrada, se procede a resolver el problema. Para esto, se ordena el arreglo de nodos según sus grados de mayor a menor. Luego, se crea un arreglo de booleanos de tamaño  $n$ , donde cada uno está inicializado en *false*. En este arreglo se guarda si el nodo ya fue visitado o no. También se crea un arreglo llamado *cids* donde se guardará la solución.

Por último, se recorre en orden el arreglo de nodos ordenados según su grado. Si un nodo no fue visitado, se agrega el nodo a *cids* y luego se marcan como visitados a todos sus adyacentes (ya que queremos que sea dominante y mínimo). Una vez que se recorre todo el arreglo, ya tenemos el *cids*, solo resta mostrarlo por pantalla.

### 4.2. Pseudocódigo

#### 4.2.1. Implementación sobre listas de adyacencia

```
Crear listas de adyacencia con el input
Crear arreglo nodos que guarda numero de nodo y grado para cada nodo

Ordenar arreglo nodos segun su grado
Crear vector de booleanos visitado de tamano n para guardar nodos visitados
Crear vector cids para guardar la solucion

Para cada nodo u en el arreglo ordenado:
    Si el nodo no fue visitado:
        agregar el nodo a cids
        Para cada nodo w en su lista de adyacencia:
            marcar w como visitado
        fin Para
    fin Si
fin Para

mostrar cids
```

#### 4.2.2. Implementación sobre matriz de adyacencia

```
Crear matriz de adyacencia con el input
Crear arreglo nodos que guarda numero de nodo y grado para cada nodo

Ordenar arreglo nodos segun su grado
Crear vector de booleanos visitado de tamano n para guardar nodos visitados
Crear vector cidm para guardar la solucion

Para cada nodo u en el arreglo ordenado:
    Si el nodo no fue visitado:
        agregar el nodo a cidm
        Para cada nodo w en su fila de la matriz de adyacencia:
            Si el nodo w es adyacente:
                marcar w como visitado
            fin Si
        fin Para
    fin Si
fin Para

mostrar cidm
```

### 4.3. Ejercicio B

Calculemos la complejidad del algoritmo. Para esto vamos a calcular la complejidad para una implementación sobre matriz de adyacencia y una sobre listas de adyacencia. A su vez dividiremos el algoritmo en tres etapas, la entrada de datos, la resolución del problema y la salida de datos.

- Matriz de adyacencia

En la entrada de datos, se crea la matriz de adyacencia, esto tiene costo temporal  $\mathcal{O}(n^2)$ , a la vez, se crea el vector de Nodos, con costo  $\mathcal{O}(n)$ . Por lo tanto, la entrada de datos tiene costo  $\mathcal{O}(n + n^2) \in \mathcal{O}(n^2)$ .

La resolución del problema, comienza ordenando el arreglo de nodos, esto tiene costo  $\mathcal{O}(n * \log(n))$ . Luego, crea los vectores con costo constante. Por último, recorre todos los nodos ( $n$ ), y por cada nodo, si no fue visitado, lo agrega a la solución con costo constante y se fija en su fila en la matriz de adyacencia quienes son sus vecinos y los marca como visitados. Esto tiene un costo de  $\mathcal{O}(n^2)$ . Por lo tanto todo el proceso tiene costo  $\mathcal{O}(n * \log(n) + n^2) \in \mathcal{O}(n^2)$ .

Por último, se muestra el *cidm* aproximado por pantalla, para eso se recorre el vector *cidm* calculado en la resolución, que tiene a lo sumo tamaño  $n$ , por lo que tiene costo  $\mathcal{O}(n)$ .

Por lo tanto, juntando las tres etapas nos queda un costo total de  $\mathcal{O}(n^2 + n^2 + n) \in \mathcal{O}(n^2)$ .

- Listas de adyacencia

Primero se deben crear las listas de adyacencia, como tengo  $m$  aristas y por cada una agrego en tiempo constante un elemento a 2 listas, esto tiene costo  $\mathcal{O}(2m) \in \mathcal{O}(m)$ .

La resolución del problema, comienza ordenando el arreglo de nodos, esto tiene costo  $\mathcal{O}(n * \log(n))$ . Luego, tenemos  $n$  listas pero la suma del tamaño de todas es de a lo sumo  $2m$ , por lo tanto, se realizarán a lo sumo  $2m$  operaciones en todo el ciclo, por lo que todo el proceso tiene costo  $\mathcal{O}(n + 2m) \in \mathcal{O}(n + m)$ .

Por último, se muestra el *cidsm* aproximado por pantalla, para eso se recorre el vector *cidsm* calculado en la resolución, que tiene a lo sumo tamaño  $n$ , por lo que tiene costo  $\mathcal{O}(n)$ .

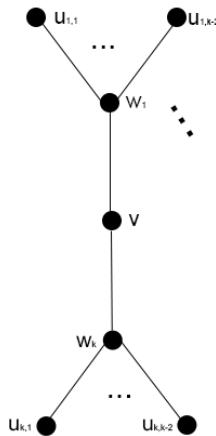
Por lo tanto, juntando las tres etapas nos queda un costo total de  $\mathcal{O}(n * \log(n) + n + m + n) \in \mathcal{O}(n * \log(n) + m)$

Entonces, la complejidad del algoritmo sobre listas de adyacencia es menor y es de:

$$\mathcal{O}(n * \log(n) + m)$$

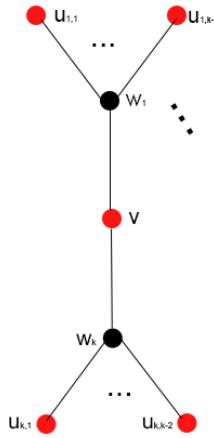
#### 4.4. Ejercicio C

Veamos un ejemplo de instancias donde la heurística golosa no funciona. Supongamos que tenemos un grafo  $G$  con un nodo central  $v$  y sean  $w_i$  con  $1 \leq i \leq k$  sus  $k$  vecinos, y supongamos que todo  $w_1$  tiene a su vez  $k - 2$  vecinos de grado 1, llamemoslos  $u_{i,j}$  con  $1 \leq i \leq k$  y  $1 \leq j \leq k - 2$ .



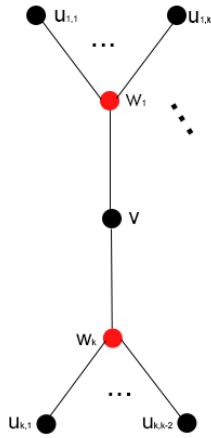
(a)

Ahora, al ordenar los nodos según su grado nos quedaría  $v$  con  $k$  vecinos, luego  $w_i$  con  $1 \leq i \leq k$ , donde cada  $w_i$  tiene grado  $k - 1$ , y por último los  $k * (k - 2)$  nodos  $u$  de grado 1. Al correr el algoritmo goloso, este arrojaría un DCIM con cardinalidad  $k * (k - 2) + 1$  ya que el algoritmo seleccionará a  $v$  y a  $u_{i,j}$  con  $1 \leq i \leq k$  y  $1 \leq j \leq k - 2$ .



(a)

Sin embargo, si seleccionamos a todos los  $w_i$  con  $1 \leq i \leq k$ , tendríamos un DCIM real con cardinalidad  $k$  que es mucho menor que el arrojado por el algoritmo.



(a)

Veamos que en este grafo  $G = (V, E)$  tenemos  $|V| = n = 1 + k * (k - 1)$  por lo tanto, los  $n$  para los cuales esta familia de instancias existe son los que satisfacen la ecuación con  $k$  entero. Ahora, el algoritmo goloso arroja una solución de tamaño  $1 + k * (k - 2)$  es decir, solo quedan sin seleccionar  $k$  nodos. Sin embargo, la solución óptima sólo utiliza  $k$  nodos.

Por lo que el error de la solución golosa es de  $1 + k * (k - 2) - k = 1 + k^2 - 2k - k = 1 + k^2 - 3k = k^2 - 3k + 1$  que es cuadrático en el tamaño de  $k$

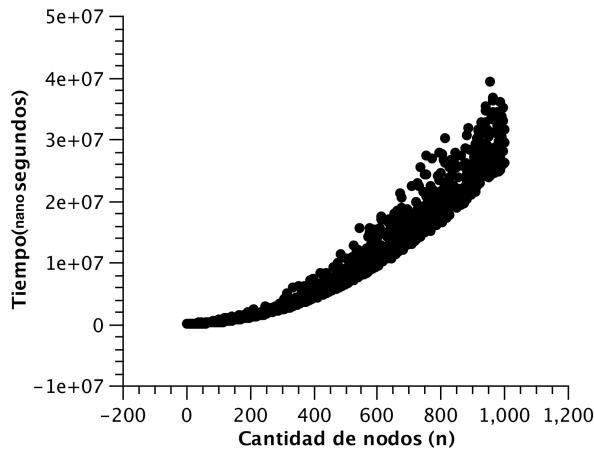
## 4.5. Ejercicio D

Para la experimentación decidimos comparar distintas familias de instancias tanto en la implementación sobre matriz como el a de listas de adyacencia y comparar la eficiencia de ambas.

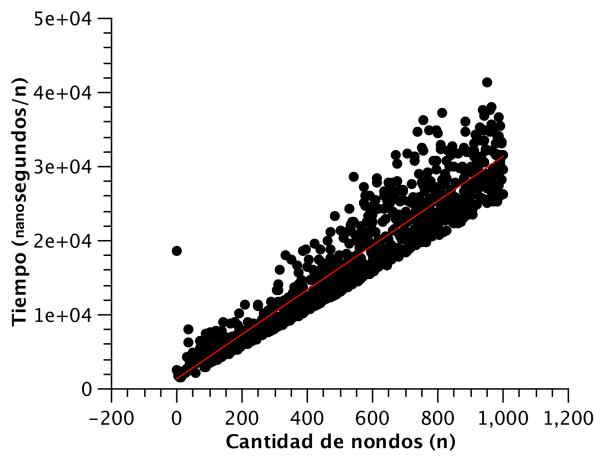
### 4.5.1. Experimentación sobre grafo completo

Se crearon instancias de grafos completos con  $1 \leq n \leq 1000$ .

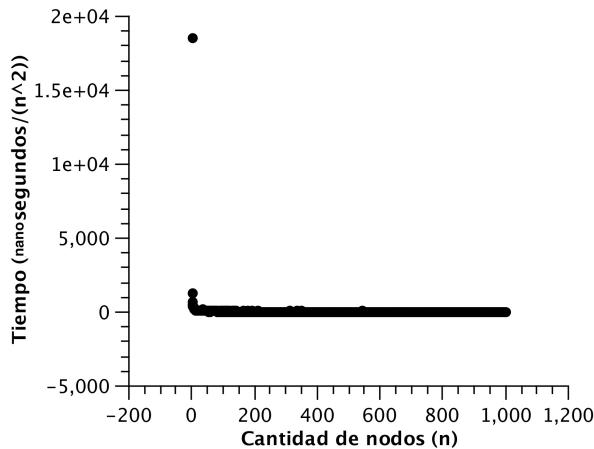
En la implementación sobre matrices arrojo los siguientes resultados:



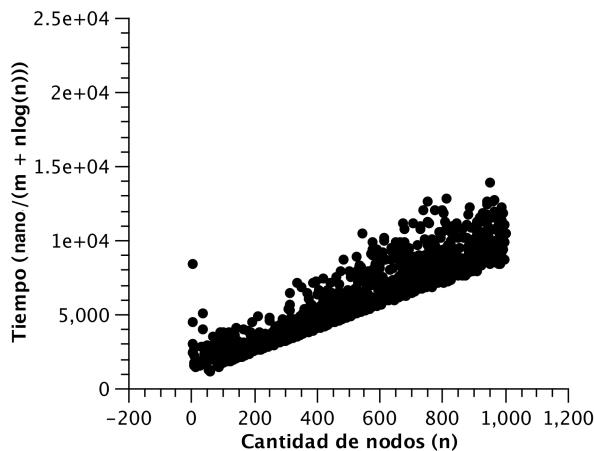
(a) Tiempos sin procesar, en milisegundos



(b) Dividiendo a los tiempos por  $n$



(a) Dividiendo a los tiempos por  $n^2$



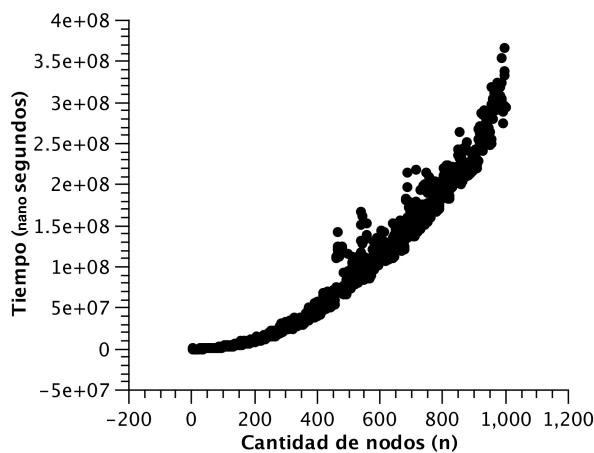
(b) Dividiendo a los tiempos por  $m + n * \log(n)$

A continuación, adjuntamos una tabla con los últimos 20 valores obtenidos en las instancias, teniendo en cuenta que los casos fueron previamente ordenados según el tamaño ( $n$ ):

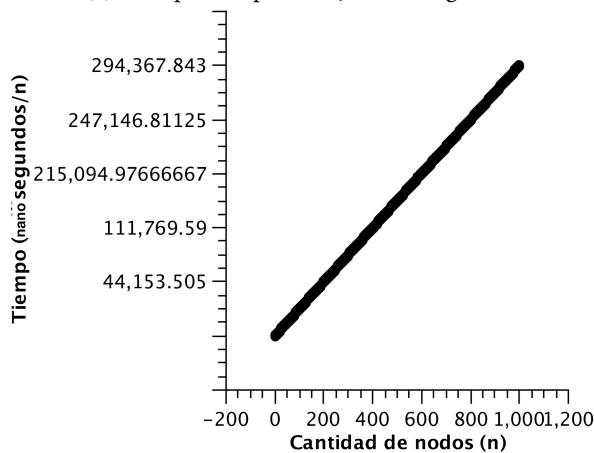
n	Tiempo(milis)	m	Tiempo(mili/(n))	Tiempo(mili/(n <sup>2</sup> )))	Tiempo(mili/(n * log(n) + m)))
980	24,845,670	479710	25,352.72448979592	25.870127030404	8,475.696536553674
981	33,412,818	480690	34,059.95718654434	34.71963015957629	11,384.934998666
982	24,857,309	481671	25,312.94195519348	25.77692663461658	8,459.892640536484
983	29,750,650	482653	30,265.15768056969	30.78856325592033	10,113.4891991821
984	25,165,002	483636	25,574.18902439025	25.99002949633155	8,544.681225102551
985	28,427,751	484620	28,860.66091370558	29.3001633641681	9,641.314760497822
986	27,111,085	485605	27,496.02941176471	27.88643956568429	9,184.088121506702
987	36,119,756	486591	36,595.49746707194	37.07750503249436	12,221.65041349897
988	33,731,556	487578	34,141.25101214575	34.5559220770706	11,400.34121094489
989	30,542,586	488566	30,882.29120323559	31.22577472521294	10,310.60678389385
990	31,176,002	489555	31,490.91111111111	31.80900112233446	10,512.26503410799
991	30,168,178	490545	30,442.15741675076	30.71862504212993	10,160.68392416765
992	25,911,896	491536	26,120.86290322581	26.33151502341311	8,717.090324045907
993	35,168,815	492528	35,416.73212487412	35.6663969031965	11,817.59489008868
994	28,113,874	493521	28,283.5754527163	28.45430126027797	9,436.079245553556
995	33,158,039	494515	33,324.66231155779	33.49212292618873	11,116.28719039287
996	26,079,861	495510	26,184.59939759036	26.28975843131563	8,733.26701997587
997	33,013,519	496506	33,112.85757271815	33.21249505789183	11,042.42206178822
998	29,605,093	497503	29,664.42184368737	29.72386958285308	9,891.007222031083
999	26,197,019	498501	26,223.24224224224	26.24949173397622	8,742.346964977023
1,000	31,613,146	499500	31,613.146	31.613146	10,537.71533333333

Como podemos ver la experimentación se condice con el cálculo teórico de la complejidad y arroja que es de  $\mathcal{O}(n^2)$ .

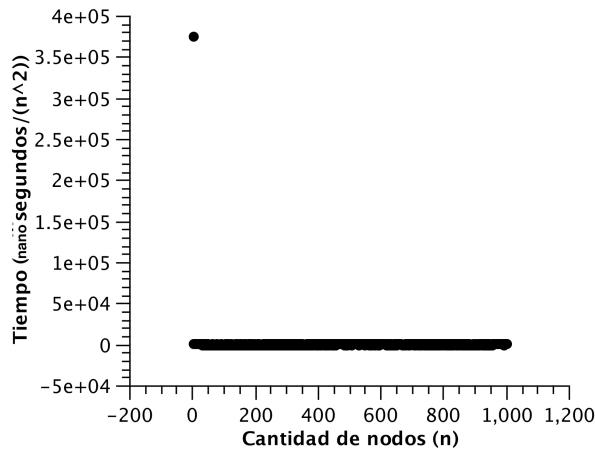
Veamos ahora los resultados en la implementación sobre listas de adyacencia:



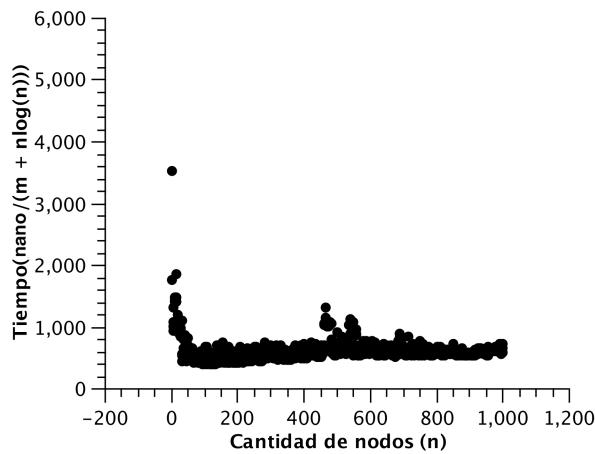
(a) Tiempos sin procesar, en milisegundos



(b) Dividiendo a los tiempos por  $n$



(a) Dividiendo a los tiempos por  $n^2$



(b) Dividiendo a los tiempos por  $m + n * \log(n)$

A continuación, adjuntamos una tabla con los últimos 20 valores obtenidos en las instancias, teniendo en cuenta que los casos fueron previamente ordenados según el tamaño ( $n$ ):

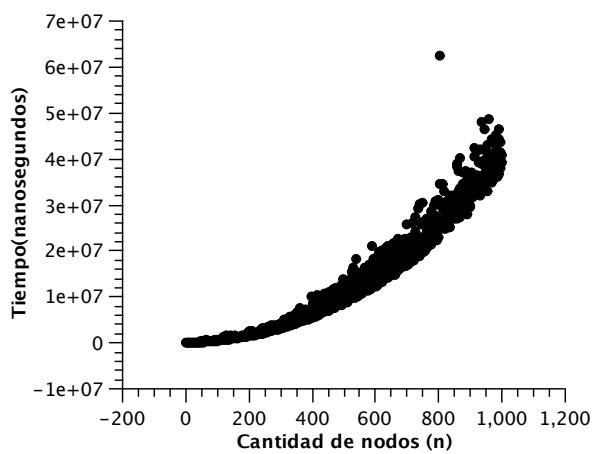
n	Tiempo(milis)	Tiempo(mili/(n))	Tiempo(mili/(n <sup>2</sup> )))	Tiempo(mili/(n * log(n) + m)))	m
980	296,879,716	302,938.48571429	309.1209037900875	615.1144826034448	479,710
981	323,033,524	329,290.03465851	335.667721364436	667.9423920520295	480,690
982	292,431,614	297,791.86761711	303.2503743555071	603.4379490268261	481,671
983	297,273,430	302,414.47609359	307.6444314278647	612.1842807953875	482,653
984	318,286,205	323,461.59044715	328.7211285032058	654.1277503491588	483,636
985	308,937,975	313,642.6142132	318.4188976783736	633.6298448756557	484,620
986	323,384,902	327,976.57403651	332.6334422276989	661.9185207263685	485,605
987	304,973,689	308,990.56636272	313.0603509247369	622.9719891480256	486,591
988	302,281,783	305,953.22165992	309.6692526922257	616.2264905731232	487,578
989	353,599,216	357,532.06875632	361.5086640609904	719.3873732062154	488,566
990	292,805,769	295,763.4030303	298.7509121518212	594.5045184459034	489,555
991	289,577,857	292,207.72653885	294.8614798575678	586.7671292958553	490,545
992	289,158,638	291,490.5625	293.841292842742	584.7394227940042	491,536
993	293,229,428	295,296.50352467	297.3781505787238	591.7801781541916	492,528
994	275,111,868	276,772.50301811	278.443161990049	554.102004459966	493,521
995	333,606,954	335,283.37085427	336.9682119138405	670.5696612605566	494,515
996	338,319,005	339,677.71586345	341.041883397042	678.679115312912	495,510
997	332,699,588	333,700.69007021	334.7048044836616	666.0709764219323	496,506
998	366,440,239	367,174.58817635	367.9104089943414	732.1539875453961	497,503
999	296,662,430	296,959.38938939	297.2566460354248	591.5530805302149	498,501
1,000	294,367,843	294,367.843	294.367843	585.8066527363184	499,500

Como podemos ver, la experimentación se condice con el cálculo teórico de la complejidad y arroja que es de  $\mathcal{O}(n * \log(n) + m)$ , aunque en un grafo completo es casi  $\mathcal{O}(n^2)$ .

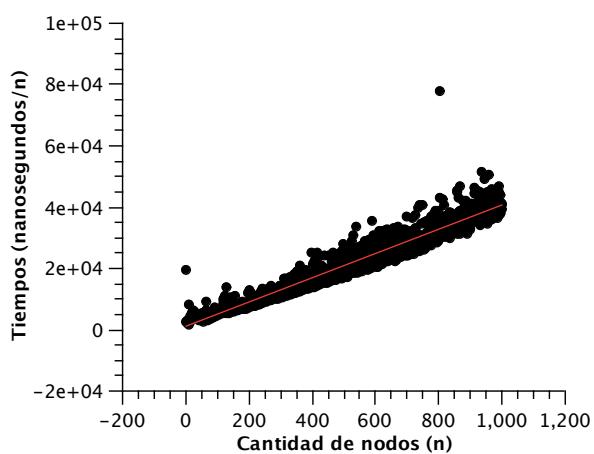
#### 4.5.2. Experimentación sobre el complemento del grafo completo

Se crearon instancias de complementos grafos completos con  $1 \leq n \leq 1000$  y  $m = 0$ .

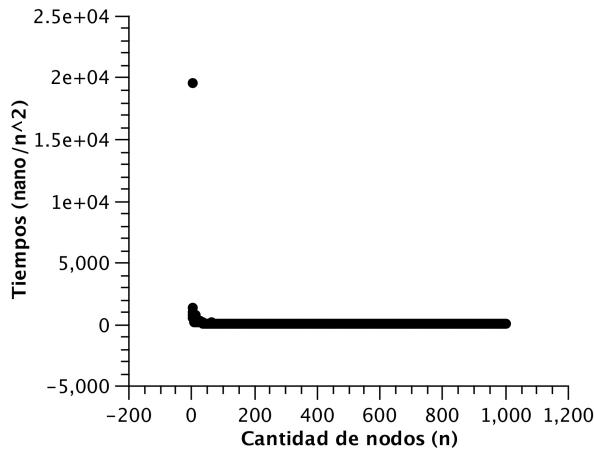
En la implementación sobre matrices arrojo los siguientes resultados:



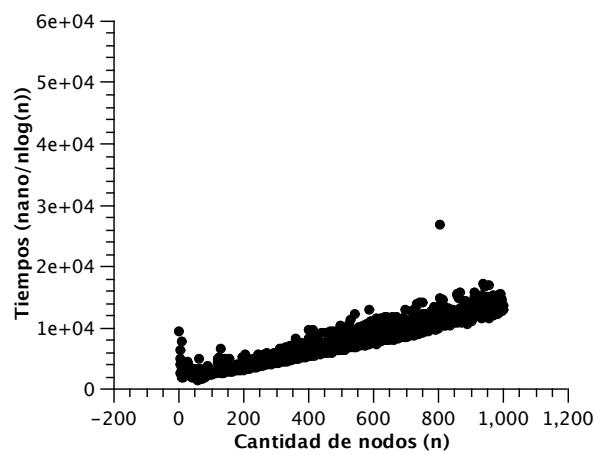
(a) Tiempos sin procesar, en milisegundos



(b) Dividiendo a los tiempos por  $n$



(a) Dividiendo a los tiempos por  $n^2$



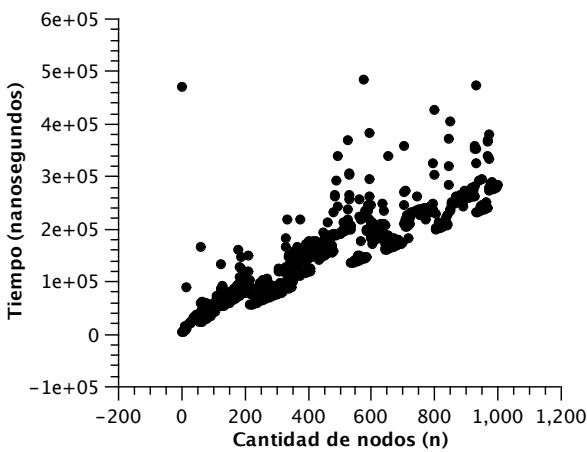
(b) Dividiendo a los tiempos por  $m + n * \log(n)$

A continuación, adjuntamos una tabla con los últimos 20 valores obtenidos en las instancias, teniendo en cuenta que los casos fueron previamente ordenados según el tamaño ( $n$ ):

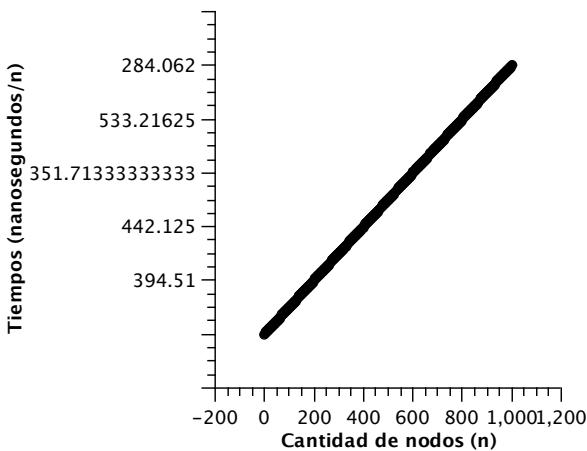
n	Tiempo(milis)	Tiempo(mili/(n))	Tiempo(mili/(n <sup>2</sup> )))	Tiempo(mili/(n * log(n) + m)))
980	41,753,937	42,606.0581632653	43.47556955435235	14,243.67703581269
981	37,549,011	38,276.25993883792	39.01759422919258	12,794.28300537819
982	39,391,112	40,113.14867617108	40.84842024050008	13,406.30148305065
983	45,239,903	46,022.28179043744	46.81819103808488	15,378.93358170479
984	36,000,279	36,585.6493902439	37.18053799821535	12,223.75853853513
985	43,702,767	44,368.29137055838	45.0439506300085	14,821.85954673998
986	40,825,343	41,405.01318458418	41.99291398030849	13,829.89827602756
987	38,862,643	39,374.5116514691	39.89312224059685	13,149.74655118415
988	36,831,350	37,278.6943319838	37.73147199593502	12,447.98660517573
989	44,397,040	44,890.83923154702	45.39013066890497	14,987.61178273532
990	39,102,722	39,497.69898989899	39.89666564636262	13,185.08310395429
991	40,118,814	40,483.16246215944	40.85081984072598	13,512.07184160979
992	46,471,583	46,846.35383064516	47.22414700669875	15,633.62968546942
993	41,389,046	41,680.81168177241	41.97463412061673	13,907.7469205387
994	43,743,891	44,007.93863179074	44.27358011246553	14,682.10400263076
995	38,823,308	39,018.4	39.21447236180904	13,015.57795408459
996	37,960,325	38,112.77610441767	38.26583946226674	12,711.63425257771
997	39,036,157	39,153.61785356068	39.27143215001071	13,056.88500714597
998	40,876,628	40,958.54509018036	41.0406263428661	13,656.80637315606
999	40,732,820	40,773.59359359359	40.81440800159519	13,593.16666151807

Como podemos ver, la experimentación se condice con el cálculo teórico de la complejidad y, aunque el no haya aristas en el grafo, la complejidad que arroja es de  $\mathcal{O}(n^2)$ .

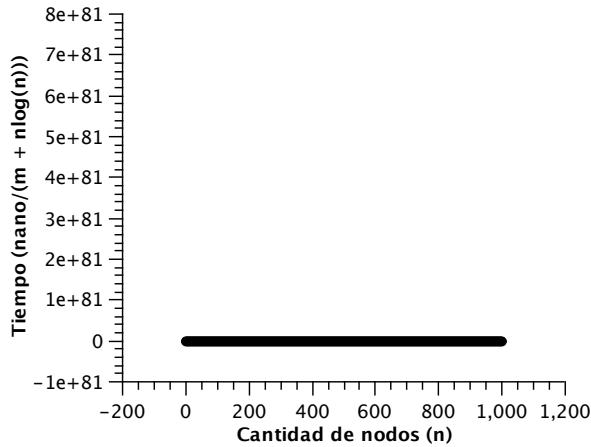
Veamos ahora los resultados en la implementación sobre listas de adyacencia:



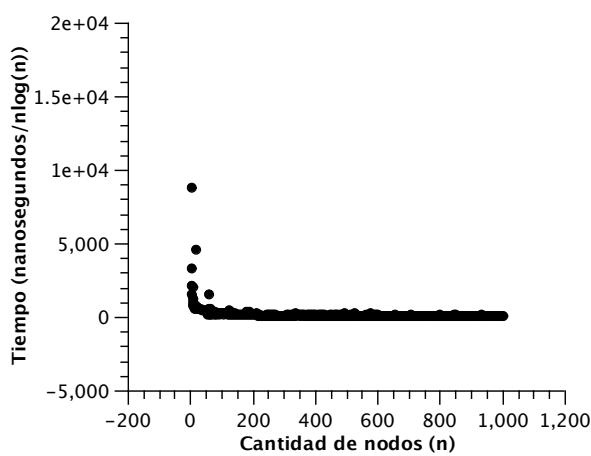
(a) Tiempos sin procesar, en milisegundos



(b) Dividiendo a los tiempos por  $n$



(a) Dividiendo a los tiempos por  $n^2$



(b) Dividiendo a los tiempos por  $m + n * \log(n)$

A continuación, adjuntamos una tabla con los últimos 20 valores obtenidos en las instancias, teniendo

en cuenta que los casos fueron previamente ordenados según el tamaño ( $n$ ):

n	Tiempo(milis)	Tiempo(mili/(n))	Tiempo(mili/(n <sup>2</sup> )))	Tiempo(mili/(n * log(n) + m)))
980	283,093	288.87040816327	0.2947657226155768	96.57257621237783
981	279,943	285.36493374108	0.2908918794506428	95.38653274714977
982	280,415	285.5549898167	0.290789195332689	95.43594581360509
983	288,893	293.88911495422	0.2989716327102968	98.20680338813817
984	276,797	281.29776422764	0.2858717116134576	93.98537417420873
985	277,190	281.41116751269	0.2856966167641526	94.00940786565884
986	276,617	280.54462474645	0.2845280169842295	93.70613178730468
987	277,663	281.3201621074	0.2850254935231977	93.95135777670718
988	281,872	285.2955465587	0.2887606746545592	95.26500875949687
989	282,107	285.24469160768	0.2884172817064555	95.23405608103857
990	277,617	280.42121212121	0.283253749617386	93.60993375526336
991	278,533	281.06256306761	0.2836150989582326	93.81029823710891
992	278,372	280.61693548387	0.2828799752861603	93.64786998548107
993	278,950	280.91641490433	0.2828966917465562	93.73412480887504
994	279,345	281.03118712274	0.2827275524373606	93.75874548093792
995	279,336	280.73969849246	0.2821504507461933	93.64785410306024
996	278,955	280.07530120482	0.2812001016112644	93.41263366232546
997	283,107	283.95887662989	0.2848133165796286	94.69414583300491
998	283,434	284.00200400802	0.2845711463006173	94.69477906957285
999	283,640	283.92392392392	0.2842081320559799	94.65501754783944
1,000	284,062	284.062	0.284062	94.68733333333333

Como podemos ver, la experimentación se condice con el cálculo teórico de la complejidad y arroja que es de  $\mathcal{O}(n * \log(n) + m)$ , sin embargo, como el grafo no tiene aristas, aquí la complejidad es de  $\mathcal{O}(n * \log(n))$  siendo así más eficiente que en la implementación sobre matriz de adyacencia.

## 5. Ejercicio 4 - Heurística de búsqueda local

*Diseñar e implementar una heurística de búsqueda local para CIDM.*

### 5.1. Ejercicio A

*Explicar detalladamente el algoritmo implementado. Plantear al menos dos vecindades distintas para la búsqueda y al menos dos soluciones iniciales.*

#### 5.1.1. Algoritmo implementado

Sea  $G = (V, E)$  un grafo simple, la heurística de búsqueda local propuesta genera una solución inicial válida, es decir un  $V' \subseteq V$  que es dominante e independiente (CID), de dos formas:

1. **Heurística constructiva golosa:** procedimiento descripto en el ejercicio anterior.
2. **Procedimiento BFS modificado:** detallado a continuación.

#### 5.1.2. Procedimiento BFS modificado:

Partimos de incluir un vértice inicial  $v$  a  $V'$  y luego vamos a ir agregando vértices a  $V'$  determinando si el vértice analizado debe incluirse en  $V'$ . El BFS modificado funciona de la siguiente manera:

Los vértices están numerados de 0 a  $n-1$ .

Creamos un vector, llamado `solucionInicial`, de tamaño  $n$  para guardar el estado de los vértices (si fue VISITADO o no)

Creamos un vector de tamaño  $n$  en donde para cada posición guardamos si pertenece al CID (INCLUIDO o no)

Al vértice inicial  $v$  lo ponemos como VISITADO y INCLUIDO y lo incluimos en la cola.

Luego mientras no esté vacía la cola:

    Sacamos el primer elemento de la cola ( $w$ ) y lo ponemos INCLUIDO.

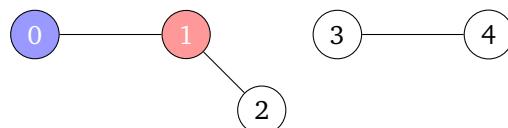
    Revisamos cada adyacente a  $w$ :

        Si algún adyacente está INCLUIDO entonces hacemos  $w = NO\ INCLUIDO$ .

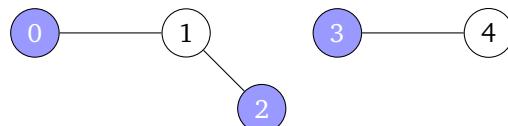
        Si el adyacente no fue VISITADO entonces lo ponemos como VISITADO y lo agregamos a la cola.

Repetimos el procedimiento para el resto de las componentes conexas, empezando por el vértice de menor numeración de la componente analizada.

A continuación mostramos un ejemplo del recorrido BFS, en donde el vértice 0 ya fue visitado y se está analizando sus adyacentes, en particular el vértice 1, el cual es provisoriamente INCLUIDO:



Para luego ser desmarcado debido a la presencia de un adyacente INCLUIDO, siendo la solución generada la siguiente:

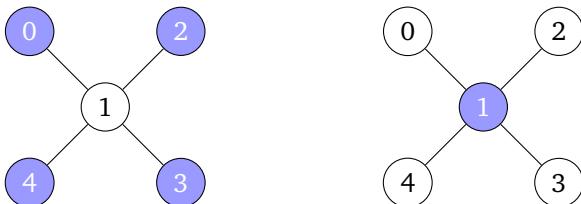


### 5.1.3. Primer Criterio de Vecindad

El primer criterio de vecindad implementado consiste en generar soluciones vecinas a partir de quitar  $k$  vértices que pertenecen al subconjunto CID de la solución inicial y agregar 1 vértice al subconjunto, donde  $k \in \mathbb{N}$  y  $k \geq 2$ . Logrando de esta manera una reducción en el cardinal del subconjunto CID de, al menos, un vértice.

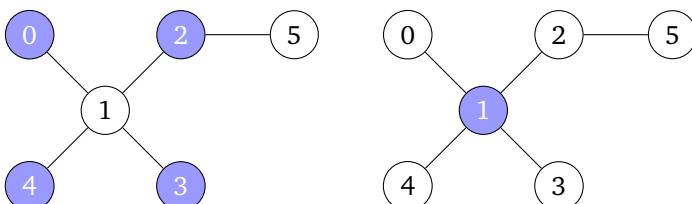
Para llevar adelante exitosamente este intercambio, debemos buscar aquellos vértices no incluidos en CID en la solución inicial, que tengan, al menos, dos vértices adyacentes incluidos en CID, para poder incluir ese vértice en la solución vecina y quitar sus adyacentes.

- Ejemplo de un cambio 4 por 1:



Sin embargo, para lograr una solución valida, los vértices quitados no pueden tener otros vértices adyacentes no incluidos en el subconjunto que, a su vez, no sean adyacentes al vértice agregado y no sean dominados por otro vértice.

- Ejemplo de solución invalida:



El procedimiento de búsqueda de los posibles soluciones vecinas funciona de la siguiente manera:

Para todo vértice,  $u$ , en el Grafo:

Creamos un vector de tamaño  $n$ , llamado `solucionAuxiliar`, al cual le copiamos el contenido de la `solucionInicial`.

Si `solucionInicial[u] = NO INCLUIDO` y  $|\text{adyacentes a } u| > 1$  entonces:

`solucionAuxiliar[u] = INCLUIDO`

`cantAdyacentesIncluidos = 0`

Para todo adyacente,  $v$ , de  $u$ :

Si `solucionInicial[v] = INCLUIDO` entonces:

`cantAdyacentesIncluidos ++`

`solucionAuxiliar[v] = NO INCLUIDO`

Si `cantAdyacentesIncluidos > 1` entonces:

Si `esSolucion?(solucionAuxiliar)` entonces:

Buscar Nuevos Vecinos a partir de la `solucionAuxiliar`

Interrumpir el ciclo

En el procedimiento descripto anteriormente falta detallar el comportamiento de la función auxiliar `esSolucion?`, la cual sera descripta en el apartado siguiente, ya que es utilizado por ambos criterios.

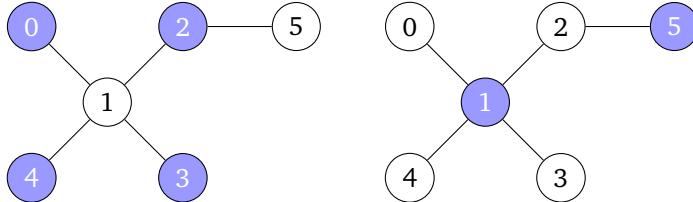
### 5.1.4. Segundo Criterio de Vecindad

El segundo criterio de vecindad implementado consiste en generar soluciones vecinas a partir de quitar  $k$  vértices que pertenecen al subconjunto CID de la solución inicial y agregar, hasta,  $k-1$  vértices al subconjunto, donde  $k \in \mathbb{N}$  y  $k \geq 2$ . Logrando de esta manera, una reducción en el cardinal del subconjunto

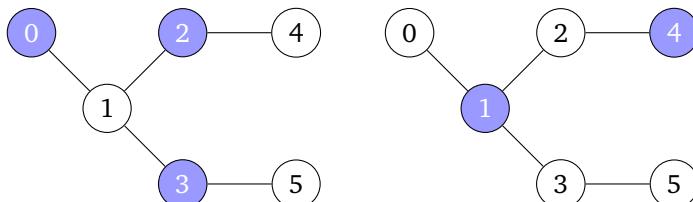
CID de, al menos, un vértice. El caso donde  $k = 2$  no difiere del criterio aplicado en la primer vecindad, ya que  $k - 1 = 1$ . Sin embargo a partir de  $k \geq 3$  se observa un comportamiento distinto, ya que podemos agregar  $k-2$  vértices para arreglar la solución, ademas del candidato original.

En este caso, para los casos no contemplados en el criterio anterior, debemos buscar vértices no incluidos en CID en la solución inicial, que tengan, al menos  $k$  vértices adyacentes incluidos en CID, donde  $k \geq 3$ , y que a su vez tengan hasta  $k-2$  vértices que son adyacentes a los adyacentes del vértice buscado que no están incluidos y no son dominados por otro vértice.

- Ejemplo de un caso 4-2, el cual fallaba en el criterio anterior:



- Caso donde falla el segundo criterio:



El procedimiento de búsqueda de los posibles soluciones vecinas funciona de la siguiente manera:

Para todo vértice,  $u$ , en el Grafo:

Creamos un vector de tamaño  $n$ , llamado `solucionAuxiliar`, al cual le copiamos el contenido de la `solucionInicial`.

Si `solucionInicial[u] = NO INCLUIDO` y  $|\text{adyacentes a } u| > 1$  entonces:

`solucionAuxiliar[u] = INCLUIDO`

`cantAdyacentesIncluidos = 0`

Para todo adyacente,  $v$ , de  $u$ :

Si `solucionInicial[v] = INCLUIDO` entonces:

`cantAdyacentesIncluidos ++`

`solucionAuxiliar[v] = NO INCLUIDO`

Si `cantAdyacentesIncluidos > 1` entonces:

`cantCambiosPosibles = cantAdyacentesIncluidos - 2`

`arreglarSolucion(solucionAuxiliar, cantCambiosPosibles)`

Si `esSolucion?(solucionAuxiliar)` entonces:

Buscar Nuevos Vecinos a partir de la `solucionAuxiliar`

Interrumpir el ciclo

FFalta detallar los procedimientos `arreglarSolucion` y `esSolucion?`, los cuales se pueden realizar en una sola función que llamaremos `esSolucion?` cuyo comportamiento es el siguiente:

- La función recibe como parámetros un vector con la solución a analizar y un entero con la cantidad de cambios posibles a realizar
- Miramos cada vértice del grafo, los cuales o están INCLUIDOS o NO INCLUIDOS en el subconjunto CID.
- Si el vértice está INCLUIDO, sus adyacentes NO pueden estar INCLUIDOS. En caso de encontrar algun adyacente INCLUIDO, sabemos que el subconjunto analizado no es solución valida.
- Si el vértice NO está INCLUIDO, entonces, al menos, 1 vértice adyacente tiene que estar INCLUIDO. En caso de no encontrar algún vértice adyacente INCLUIDO, tenemos dos casos:

1. La variable entera que representa la cantidad de cambios posibles es 0. En este caso sabemos que el subconjunto analizado no es solución valida.
2. La variable entera que representa la cantidad de cambios posibles es mayor a 0. En este caso el vértice pasa a estar INCLUIDO en el subconjunto, manteniéndose la validez de la solución, ya que el vértice NO tiene adyacentes INCLUIDOS. También reducimos la cantidad de cambios posibles en una unidad.

■ En pseudocódigo:

```
Como entrada tenemos el vector solucionAuxiliar con la solución a analizar y el entero cantCambiosPosibles, que tiene la cantidad de vértices que podemos incluir.
```

Creamos una variable booleana, esSolucion inicializada en true.

Luego, para todo vértice, u, en el Grafo:

```
Si solucionAuxiliar[u] = INCLUIDO y |adyacentes a u| > 0 entonces:
```

```
    Para todo adyacente, v, de u:
```

```
        Si solucionInicial[v] = INCLUIDO entonces:
```

```
            esSolucion = false
```

```
            Interrumpir el ciclo
```

```
Sino Si |adyacentes a u| > 0, entonces:
```

```
    adyacenteIncluido = false
```

```
    Para todo adyacente, v, a u:
```

```
        Si solucionInicial[v] = INCLUIDO entonces:
```

```
            adyacenteIncluido = true
```

```
    Si not(adyacenteIncluido) y cantCambiosPosibles = 0 entonces:
```

```
        esSolucion = false
```

```
        Interrumpir el ciclo
```

```
    Sino Si not(adyacenteIncluido) entonces:
```

```
        solucionAuxiliar[u] = INCLUIDO
```

```
        cantCambiosPosibles --
```

```
Sino entonces:
```

```
    esSolucion = (solucionAuxiliar[u] = INCLUIDO)
```

Es necesario aclarar que para el primer criterio de vecindad, la cantidad de cambios posibles es 0.

## 5.2. Ejercicio B

*Calcular el orden de complejidad temporal de peor caso de una iteración del algoritmo.*

La estructura de datos que utilizamos para representar los grafos son vectores con listas, donde cada posición del vector representa un vértice y las listas contienen los adyacentes a ese vértice.

A partir de los procedimientos expuestos en el punto anterior, pasamos a analizar la complejidad de la heurística propuesta, para solo una iteración:

### 1. Solución Inicial

- Heurística Golosa:  $\mathcal{O}(n * \log(n) + m)$ . Justificada anteriormente.
- BFS modificado: Las cambios implementados en el BFS no alteran su complejidad original, siendo la misma  $\mathcal{O}(n + m)$ .<sup>5</sup>

### 2. Primer Criterio de Vecindad

Tenemos un ciclo que se repite  $n$  veces, el cual tiene varias operaciones que se realizan internamente:

<sup>5</sup>Referencia [https://en.wikipedia.org/wiki/Breadth-first\\_search](https://en.wikipedia.org/wiki/Breadth-first_search)

- Creación de un vector tamaño  $n$  y copia de contenido:  $\Theta(n)$
- Comparaciones y Asignaciones:  $\mathcal{O}(1)$
- Ciclo de los adyacentes, cuya complejidad, sumada a la del ciclo principal, es  $\mathcal{O}(n + m)$ .
- Complejidad de la función esSolucion?:  $\mathcal{O}(n + m)$ . Detallada en el punto 4.

Por lo tanto la complejidad total de este procedimiento es:  $\mathcal{O}(n * (n + n + m) + n + m)$ , lo cual es:  $\mathcal{O}(n * (n + m))$

### 3. Segundo Criterio de Vecindad

Misma situación que el punto anterior, tenemos un ciclo que se repite  $n$  veces, el cual tiene varias operaciones que se realizan internamente:

- Creación de un vector tamaño  $n$  y copia de contenido:  $\Theta(n)$
- Comparaciones y Asignaciones:  $\mathcal{O}(1)$
- Ciclo de los adyacentes, cuya complejidad, sumada a la del ciclo principal, es  $\mathcal{O}(n + m)$ .
- Complejidad de la función esSolucion?:  $\mathcal{O}(n + m)$ . Detallada en el punto 4.

Por lo tanto la complejidad total de este procedimiento es:  $\mathcal{O}(n * (n + n + m) + n + m)$ , lo cual es:  $\mathcal{O}(n * (n + m))$

### 4. Procedimiento esSolucion?

Tenemos un ciclo que se repite  $n$  veces, el cual tiene varias operaciones que se realizan internamente:

- Comparaciones y Asignaciones:  $\mathcal{O}(1)$
- Ciclo de los adyacentes, cuya complejidad, sumada a la del ciclo principal, es  $\mathcal{O}(n + m)$ .

Por lo tanto la complejidad total de este procedimiento es:  $\mathcal{O}(n + m)$ .

Podemos concluir que la complejidad temporal de la heurística es independiente del procedimiento utilizado para armar la solución inicial, y que utilizando el primer o segundo criterio de vecindad la complejidad es la misma:  $\mathcal{O}(n * (n + m))$ .

#### Cota superior para la cantidad de iteraciones:

Sabemos que cada iteración de las vecindades reduce, como mínimo, en 1 el cardinal del subconjunto CID. Por lo tanto, si partimos de una solución inicial en donde el cardinal del subconjunto es asintóticamente igual a la cantidad de vértices del Grafo, es posible que iteraremos hasta  $n - 1$  veces, hasta alcanzar una solución de 1 vértice. Es evidente que este es un caso extremo, de difícil realización, sin embargo brinda una cota superior a la cantidad de iteraciones.

## 5.3. Ejercicio C

*Realizar una experimentación que permita observar la performance del algoritmo comparando los tiempos de ejecución y la calidad de las soluciones obtenidas.*

Para llevar adelante las comparaciones, generamos instancias aleatorias con las siguientes características:

1. Tomamos 10 muestras por cada valor de la cantidad de nodos  $n$ , desde 5 hasta 30, promediando sus valores.
2. Para cada valor de  $n$ , generamos instancias para cada valor posible de  $m$ , es decir valores que van desde 0 hasta  $\frac{n*(n-1)}{2}$ , asegurándonos la validez de los aristas generadas, es decir que no haya ejes repetidos ni loops.

Contamos con 4 variantes posibles, que combinan dos modos de generar la solución inicial (BFS y Golosa) y dos criterios de vecindad:

- **Tipo B1:** Utilizamos el BFS modificado para la solución inicial, y el primer criterio de vecindad.
- **Tipo G1:** Utilizamos la heurística golosa para la solución inicial, y el primer criterio de vecindad.
- **Tipo B2:** Utilizamos el BFS modificado para la solución inicial, y el segundo criterio de vecindad.
- **Tipo G2:** Utilizamos la heurística golosa para la solución inicial, y el segundo criterio de vecindad.

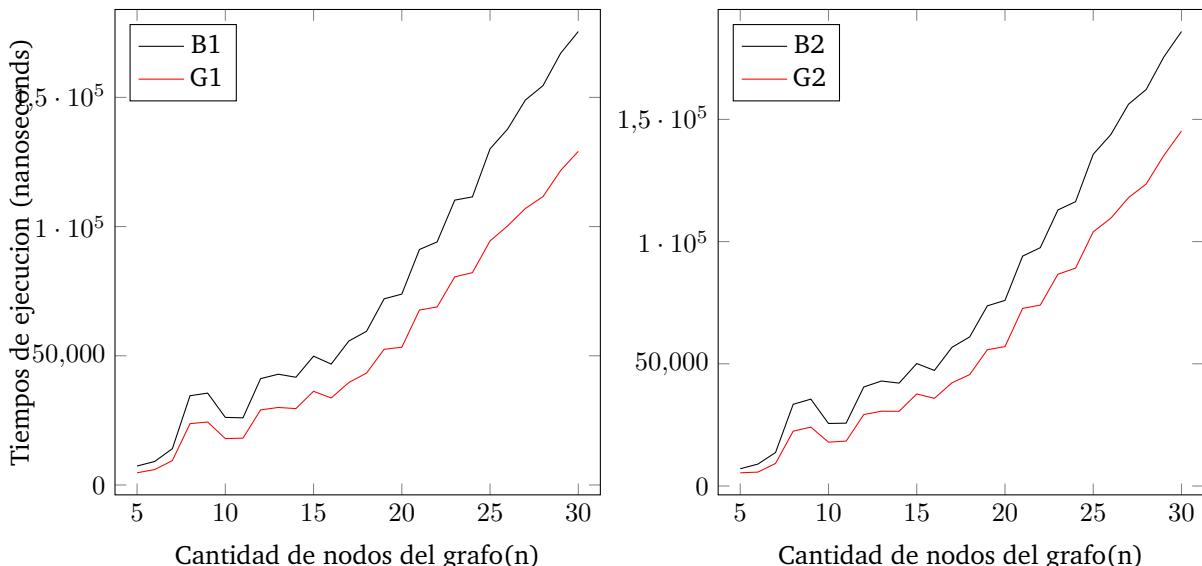
Es importante destacar los siguientes puntos:

- En el caso de las comparaciones de los tiempos de ejecución, el eje de las abscisas esta determinado por la cantidad de nodos del grafo y el eje de las ordenadas representará el tiempo que tardo el algoritmo en nanosegundos.
- En el caso de las comparaciones de las soluciones obtenidas, el eje de las abscisas esta determinado por la cantidad de nodos del grafo y el eje de las ordenadas representará el tamaño de la solución obtenida.
- La toma de tiempos en nanosegundos se realizó usando la librería 'chrono' de C++.
- La aleatorización de los ejes en sus extremos particulares se realizó con la función std::rand.

### 5.3.1. Tiempos de ejecución

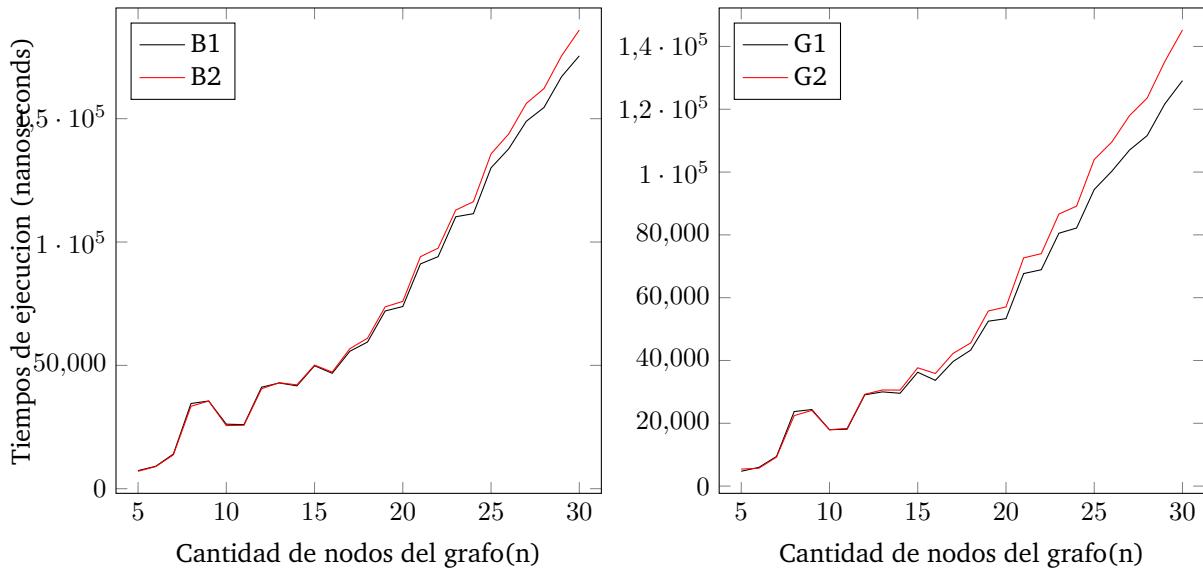
Primero comparamos las maneras de obtener la solución inicial, para luego comparar las vecindades entre si. Por último comparamos las 4 combinaciones entre si.

- **Comparación BFS contra Goloso**



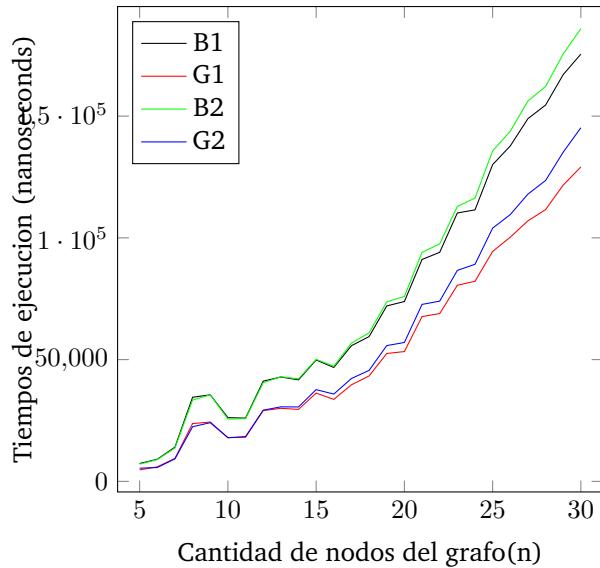
Los graficos ponen en evidencia la paridad en términos de ejecución temporal de ambas variantes. Si bien existe una mínima diferencia en favor del Goloso, no es suficiente para concluir que esta es la mejor en términos de ejecución temporal.

- **Comparación Primer Criterio de Vecindad contra Segundo Criterio de Vecindad**



En este caso la similitud es mas evidente, por lo cual no podemos concluir que una sea mejor que la otra.

■ Comparación B1-G1-B2-G2

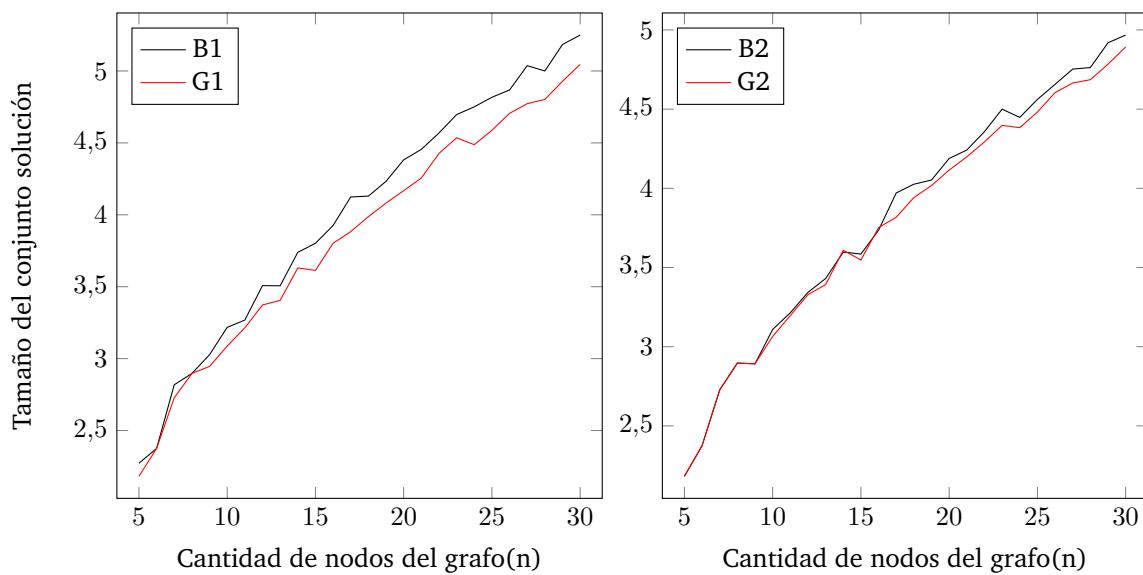


La ultima comparación pone en evidencia que la perfomance en terminos de ejecución temporal no difiere para las distintas variantes, lo cual va en sintonia con el hecho que todas tienen la misma complejidad temporal teorica, mostrada en el punto anterior. Por lo tanto la elección de que combinación es mejor sera resuelto por la calidad de las soluciones obtenidas con cada una de ellas.

### 5.3.2. Calidad de la solución

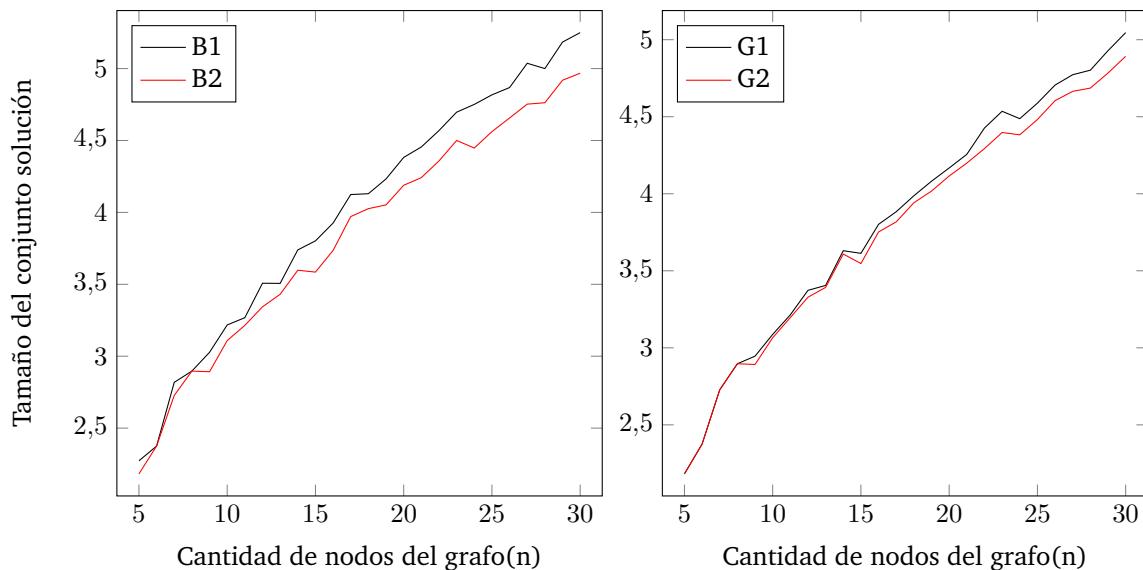
Repetimos el mismo procedimiento usado en la comparación de tiempos de ejecución, sin dejar de mencionar el hecho de que al generar las instancias de la manera descripta anteriormente, nos encontramos con el hecho que el tamaño de las soluciones no son enteros, lo cual responde a que promediamos para cada  $n$  el tamaño de la solución para todos los valores de  $m$  posibles.

■ Comparación BFS contra Goloso



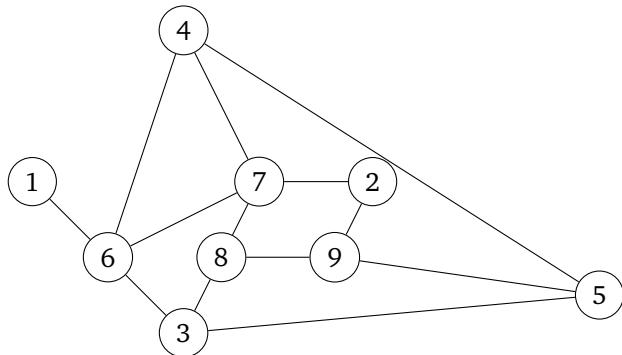
En ambos casos, se aprecia una paridad entre ambos soluciones, sin embargo hay una leve tendencia hacia la heurística golosa como mejor procedimiento para construir la solución inicial.

■ Comparación Primer Criterio de Vecindad contra Segundo Criterio de Vecindad

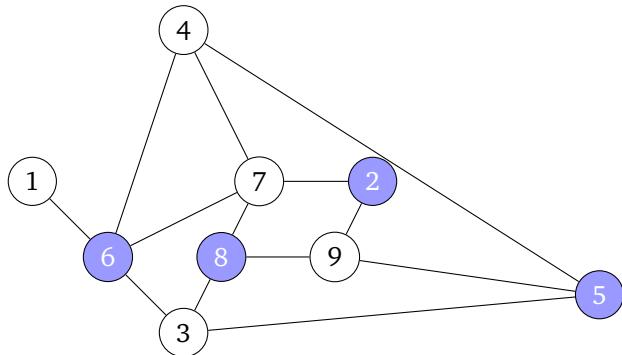


Estos resultados nos inclinan a pensar que el segundo criterio de vecindad es el mejor criterio, lo cual es coherente con el hecho que el segundo criterio de vecindad es capaz de arreglar soluciones que el primer criterio de vecindad da como invalidas. Es decir, podríamos argumentar que como el primer criterio de vecindad es un caso particular del segundo criterio de vecindad (cuando no incluimos ning n v rtice al subconjunto, mas all  del candidato) deber a ser siempre mejor al primero. Sin embargo, la experimentaci n muestra casos donde el primer criterio es mejor, y eso se da en instancias donde arreglar una soluci n invalida nos inhibe de seguir explorando la soluci n original en b squeda de un mejor caso. Por ejemplo:

■ Grafo de  $n = 9$  y  $m = 13$

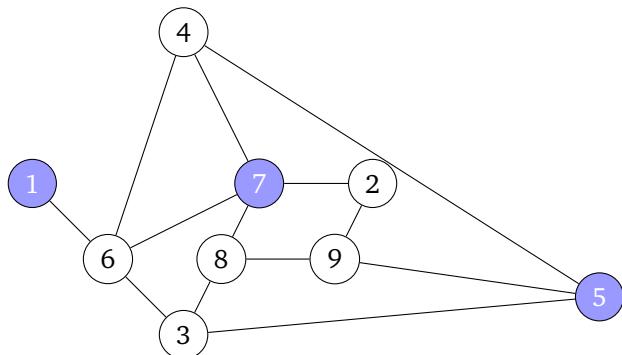


- **Solución Golosa:** como el mayor grado, 3, es compartido por varios vértices, empezamos por el menor numero de etiqueta, que es el 5, y asi sucesivamente.

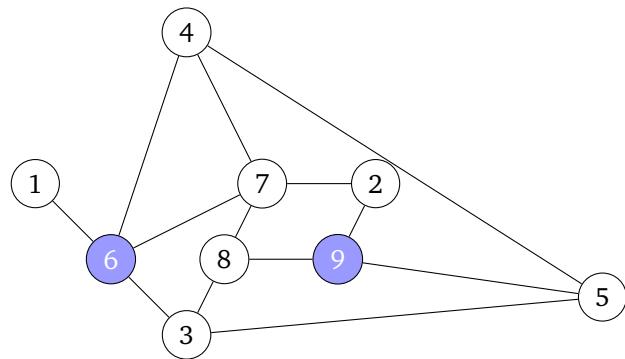


- **Solución aplicando el segundo criterio de vecindad a la solución golosa:** tenemos 4 candidatos que cumplen con el hecho de tener 2 o mas adyacentes incluidos: 3, 4, 7, 9.

1. Vértice 3: incluimos el 3, quitamos el 5, 6 y 8, por lo tanto podemos incluir un vértice mas, que sera el 1. Sin embargo el 4 no es dominado por nadie, por lo cual no es una solución valida.
2. Vértice 4: incluimos el 4, y quitamos el 5 y el 6. No podemos incluir ningun vértice mas, por lo tanto el 1 no es dominado por nadie, no es una solución valida.
3. Vértice 7: incluimos el 7 y quitamos el 2, 6 y 8. Incluimos el 1, y nos queda una solución valida.
4. Vértice 9: no es contemplado, ya que obtuvimos una solución mejor con el vértice 7. Solución que ya no es posible mejorar.

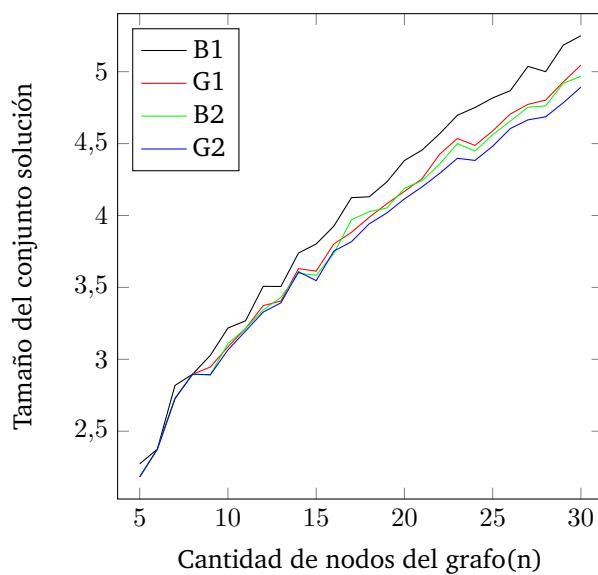


- **Solución aplicando el primer criterio de vecindad a la solución golosa:** los candidatos son los mismos que en el segundo criterio, sin embargo al no poder incluir vértice mas allá del candidato, la solución probando con el vértice 7 es invalida, ya que el 1 no es dominado por nadie. Por lo tanto probamos con incluir el 9, y quitar el 2, 5 y 8. Al hacer esto nos queda una solución valida de menor cardinal que en la solución golosa original y que si hubiésemos utilizado el segundo criterio



Sin embargo estos tipos de instancias son muy particulares, por lo cual podemos afirmar que la **mejor combinación para la heurística de búsqueda local** planteada es aquella que toma como **solución inicial** la generada por la **heurística golosa** y luego es mejorada por el **segundo criterio de vecindad**.

■ Comparación B1-G1-B2-G2



## 6. Ejercicio 5 - Metaheurística GRASP

*Diseñar e implementar un algoritmo para CIDM que use la metaheurística GRASP.*

### 6.1. Ejercicio A

*Explicar detalladamente el algoritmo implementado. Plantear distintos criterios de parada y de selección de la lista de candidatos (RCL) de la heurística golosa aleatorizada.*

#### 6.1.1. Idea general

Como pide el enunciado, la idea general del algoritmo es usar la metaheurística GRASP, para lo cual es necesario tener implementaciones de: heurística constructiva golosa y heurística de búsqueda local, que fueron convenientemente implementadas en los puntos anteriores.

La estructura general de un algoritmo GRASP es:

1. Poner en `mejor_solucion` una primera solución `Random`.
2. Mientras no se cumpla el criterio de parada hacer:
  3. Poner en `nueva_solucion` una solución usando la función `ConstruirGreedyRandom()`
  4. Intentar mejorar la `nueva_solucion` usando la función `BúsquedaLocal()`
  5. Si `costo(nueva_solucion) < costo(mejor_solucion)` hacer:
    6. Poner en `mejor_solucion` la `nueva_solucion`

En nuestro algoritmo se implementó de la siguiente forma:

1. Se utilizó para la primera 'mejor\_solución' `Random` el mismo método `ConstruirGreedyRandom` que se utiliza al generar una solución golosa randomizada.
2. Los criterios de parada considerados se detallan más adelante.
3. La función `ConstruirGreedyRandom` es una variación del algoritmo goloso implementado para el Ejercicio 3 (agregando lista de candidatos), detallado más adelante.
4. La función `BúsquedaLocal` es idéntica al algoritmo implementado para el Ejercicio 4. En la experimentación se probó con los criterios de vecindad 1 y 2 expuestos en ese mismo Ejercicio.
5. Definimos el costo de una solución como la cantidad de nodos de dicha solución, por lo que decimos que una es mejor que otra si la primera tiene menor cantidad de nodos.
6. Si encontramos una solución con menor cantidad de nodos que la mejor hasta ese momento, la guardamos como mejor solución.

#### 6.1.2. Criterios de parada

Los criterios de parada que se utilizaron para la implementación se pensaron en función de la cantidad de nodos del grafo original:

1. Criterio 1: realizar  $n$  iteraciones, con  $n$  la cantidad de nodos del grafo.
2. Criterio 2: sea  $k$  una constante, seguir iterando hasta que la mejor solución parcial no se mejore durante  $k$  ciclos seguidos.

#### 6.1.3. Selección de lista de candidatos (RCL)

Al algoritmo con heurística constructiva golosa del Ejercicio 3 se lo modificó de la siguiente forma:

- En vez de iterar en los elementos del array de nodos ordenados por grado, iteramos hasta que hayamos visitado  $n$  nodos usando un contador, ya que no necesariamente vamos a visitar secuencialmente todos los nodos desde el índice 0 hasta el  $(n-1)$ -ésimo.

- Dentro del ciclo principal, lo primero que hacemos es elegir el índice de un nodo para agregar a la solución.

A diferencia del algoritmo goloso original, que elegíamos siempre el nodo con grado más alto no visitado hasta ese momento, ahora vamos a tener una lista de candidatos (nodos) a agregar a la solución, y de todos ellos vamos a elegir alguno de manera aleatoria.

La lista de candidatos se construye de dos formas posibles:

1. Criterio 1: tomando como referencia el nodo con grado más alto no visitado hasta ese momento. Si  $d_{max}$  es dicho grado, agregaremos a la lista de candidatos todos los nodos que tengan grado al menos  $\alpha * d_{max}$  ( $\alpha$  constante). Es decir, sea  $d$  el grado del nodo, los consideramos si  $d \geq d_{max} * \alpha$ . Donde  $\alpha$  es un valor entre 0 y 1.
2. Criterio 2: en vez de tomar como referencia el grado de mayor elemento, simplemente tomamos los  $k$  elementos de mayor grado no visitados hasta el momento, con  $k$  un valor constante entero.

Esto nos asegura que, si bien el nodo a agregar a la solución es aleatorio, se encuentra dentro de cierto grupo de nodos mejores que otros.

- Luego de que se eligió un nodo, se lo agrega a la solución, y luego se lo borra de los nodos posibles para futuras iteraciones (se lo marca como *visitado*). Además, se aumenta en uno la cantidad de nodos visitados.
- Por último, se itera sobre todos los nodos adyacentes al elegido, borrándolos de los nodos posibles y aumentando en uno el contador de nodos visitados.

#### 6.1.4. Pseudocódigo

El esquema general del algoritmo GRASP ya se mostró en la sección Idea General, y el algoritmo de Búsqueda Local es identico al utilizado en el Ejercicio 4, por lo que mostraremos aquí solo el pseudocódigo de la función ConstruirGreedyRandom (con el primer criterio de lista de candidatos, el segundo simplemente elige los  $k$  nodos de mayor grado, con  $k$  una constante):

```

Poner nodos = un vector de structs Nodo, que tiene el indice del nodo y su grado
en el grafo, de tamaño n.
Ordenar dicho conjunto de mayor a menor grado.
Poner solucion = un vector de enteros inicializados en 0. El valor en cada indice
representa si el nodo con dicho indice pertenece o no a la solución.
Poner nodos_visitados = 0
Mientras nodos_visitados < n hacer:
    Poner mejor_grado = nodos[0].grado
    Poner limite_indice = 0
    Para i desde 0 hasta |nodos| hacer:
        Si nodos[i].grado >= mejor_grado - mejor_grado * alpha hacer:
            limite_indice = i
        Sino
            Salir ciclo Para
        Fin Si
    Fin Para

    Poner indice_nuevo = random_in_range(0, min(limite_indice, |nodos|-1))
    Poner nodo_nuevo = nodos[indice_nuevo].indice
    Agrego nodo_nuevo al vector solucion y lo borro del vector nodos
    Incrementar nodos_visitados en uno

    Para v en Adyacentes(nodo_nuevo) hacer:
        Si v esta en nodos hacer:
            Borrar v del vector nodos
            Incrementar nodos_visitados en uno
        Fin Si
    Fin Para
Fin Mientras
Devolver solucion

```

## 6.2. Ejercicio B

*Realizar una experimentación que permita observar los tiempos de ejecución y la calidad de las soluciones obtenidas.*

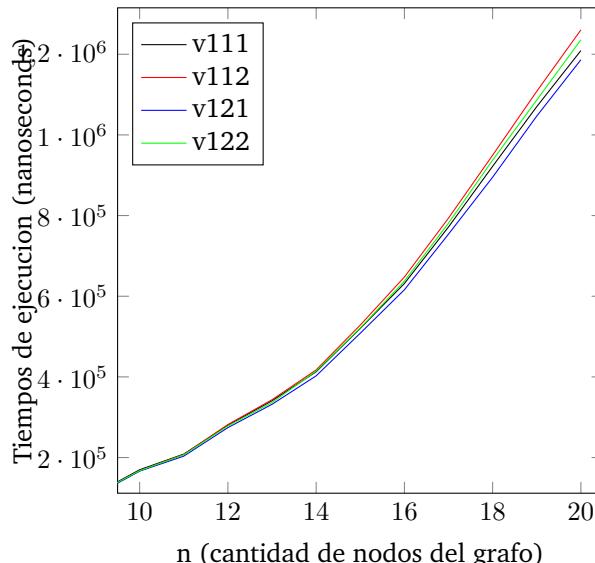
Consideraciones:

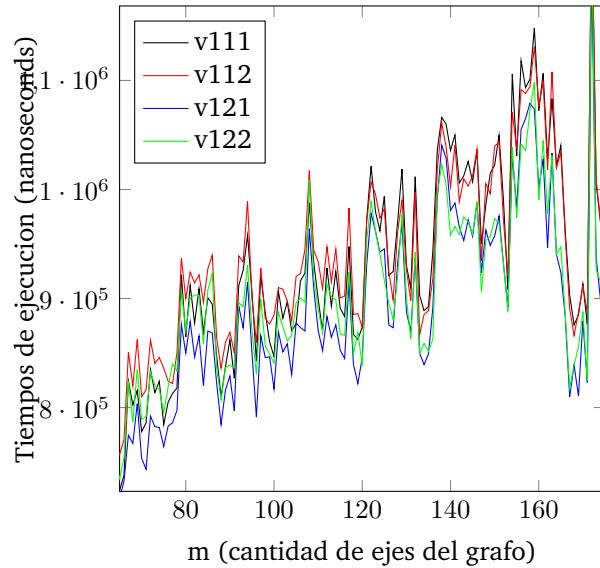
- Todas las tomas de tiempos fueron hechas bajo las mismas condiciones (computadora, fuente de energía, mínima cantidad de procesos abiertos).
- Varias muestras fueron tomadas de cada dato para luego ser promediados y así dar datos más confiables.
- En base a lo explicado anteriormente, realizamos experimentos basando los criterios de los algoritmos para así crear diferentes versiones del mismo.
- La diferencia entre ellas son los siguientes criterios:
  - Criterio de parada 1: realizar  $n$  iteraciones, con  $n$  la cantidad de nodos del grafo.
  - Criterio de parada 2: sea  $k$  una constante, seguir iterando hasta que la mejor solución parcial no se mejore durante  $k$  ciclos seguidos.
  - Criterio de lista de candidatos aleatoria 1: tomamos los nodos que tengan grado al menos  $\alpha * d_{max}$  ( $\alpha$  constante) con  $d_{max}$  el mayor de los grados entre los nodos no visitados.
  - Criterio de lista de candidatos aleatoria 2: tomar los  $k$  elementos de mayor grado no visitados hasta el momento, con  $k$  un valor constante entero.

- Criterio de vecindad en búsqueda local 1: generar soluciones vecinas a partir de quitar k vértices que pertenecen al subconjunto CID de la solución inicial y agregar 1 vértice al subconjunto.
- Criterio de vecindad en búsqueda local 2: generar soluciones vecinas a partir de quitar k vértices que pertenecen al subconjunto CID de la solución inicial y agregar, hasta, k-1 vértices al subconjunto.
- Las versiones sobre las cuales experimentamos fueron:
  - v111: criterio parada 1, criterio lista candidatos 1, criterio búsqueda local 1.
  - v112: criterio parada 1, criterio lista candidatos 1, criterio búsqueda local 2.
  - v121: criterio parada 1, criterio lista candidatos 2, criterio búsqueda local 1.
  - v122: criterio parada 1, criterio lista candidatos 2, criterio búsqueda local 2.
  - v211: criterio parada 2, criterio lista candidatos 1, criterio búsqueda local 1.
  - v212: criterio parada 2, criterio lista candidatos 1, criterio búsqueda local 2.
  - v221: criterio parada 2, criterio lista candidatos 2, criterio búsqueda local 1.
  - v222: criterio parada 2, criterio lista candidatos 2, criterio búsqueda local 2.
- Cuando comparamos los tiempos de ejecución, el eje 'x' determina el tamaño de la entrada (que puede ser la cantidad de nodos, ejes del grafo o la suma de ellos). El eje 'y' representará el tiempo que tardo el algoritmo en nanosegundos.
- La toma de tiempos en nanosegundos se realiza con la librería 'chrono' de C++.
- La aleatorización de algunas variables (por ejemplo, para elegir un candidato de la Lista Restringida de Candidatos RCL) se hizo con la función std::rand.
- Para la comparacion de calidad de resultados, tomamos como eje 'x' la cantidad de nodos del grafo, le agregamos aristas aleatoriamente y luego corremos el algoritmo con las diferentes versiones. El eje 'y' en este caso seria la cantidad de nodos del CIDM que devuelve la version del algoritmo. Finalmente promediamos los resultados para cada 'x'.

### 6.2.1. Tiempos de ejecución

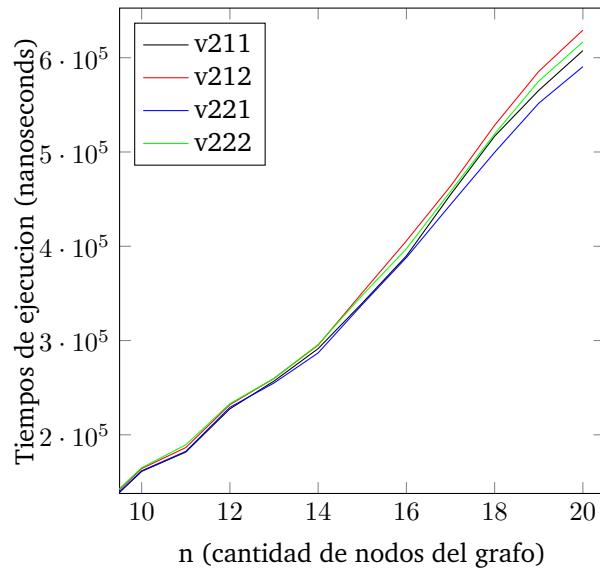
Primero, vamos a comparar de a cuatro versiones a la vez (todas al mismo tiempo es mas dificil de ver la diferencia). Como dijimos antes, los ejes de los grafos fueron puestos de forma aleatoria. Por cada instancia construida, tomamos los tiempos con cada versión muchas veces y promediamos los tiempos resultantes:

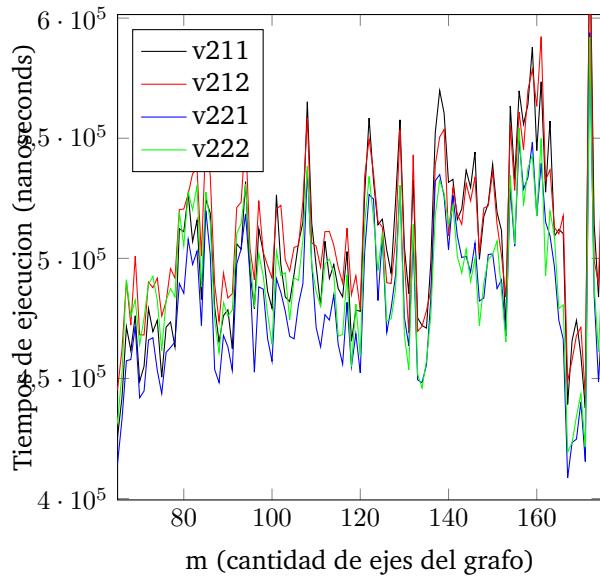




Vemos que la versión v121 fue la que menos tiempo tardó en ejecutarse. Si bien la comparación bajo la cantidad de aristas parece inestable, esto era predecible ya que los ejes son puestos de forma aleatoria y GRASP actúa en varias fases. Como la posición de los ejes es clave para que los algoritmos encuentren CIDMs, los tiempos del gráfico varian. Mas alla de eso, es importante ver que varian de forma parecida entre las distintas versiones y que además, se mantienen las diferencias entre versiones de cuando comparamos por cantidad de nodos (es decir, la que fue más rápida en una, también lo fue en la otra).

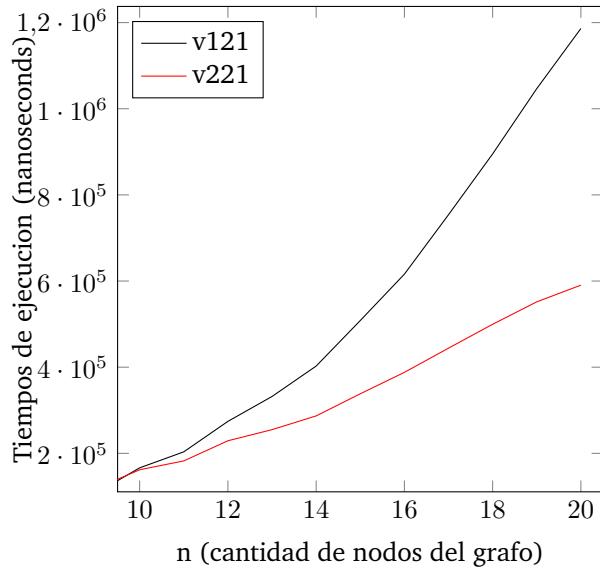
Ahora veamos las otras cuatro versiones:

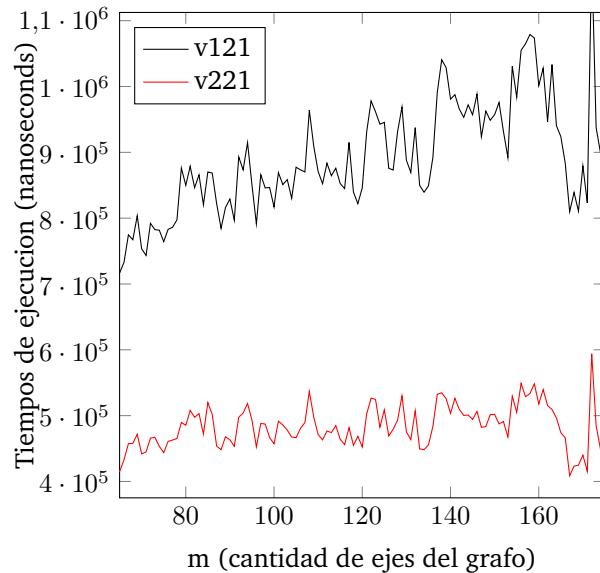




En este experimento, la versión v221 fue la que tardó menos. Hasta ahora observamos que el segundo criterio en la lista de candidatos y el primer criterio en la búsqueda local fueron los más veloces.

Comparemos las versiones v121 con la versión v221:

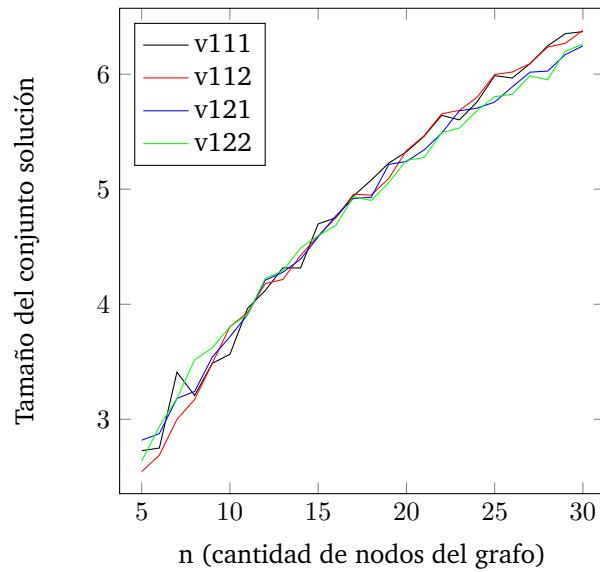




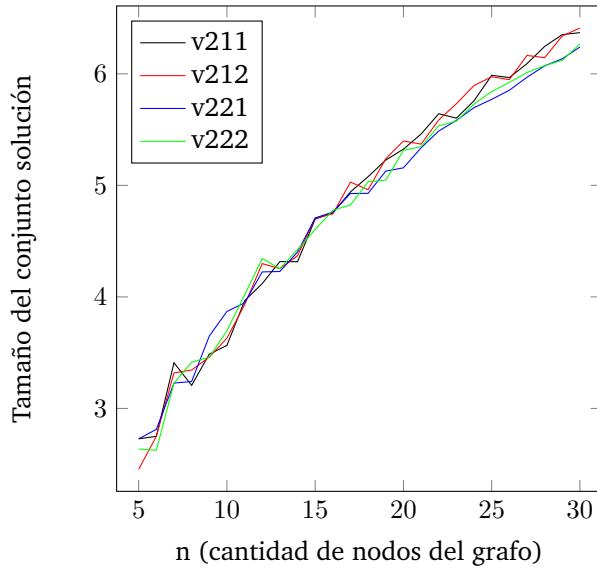
Viendo estos últimos gráficos es notoria la diferencia. La versión v221 es enormemente más rápida que la versión v121. Entonces concluimos que la gran diferencia reside en el criterio de parada de las iteraciones de GRASP. De por si el criterio 2 es más lógico, ya que el otro depende de la cantidad de nodos y este va más a la raíz del problema, es decir, cuánto se puede optimizar la función hasta que ya no valga la pena. Vale preguntarse, entonces, cuál es el criterio que da los resultados más precisos. Esa es una pregunta que responderemos a continuación.

### 6.2.2. Calidad de las respuestas

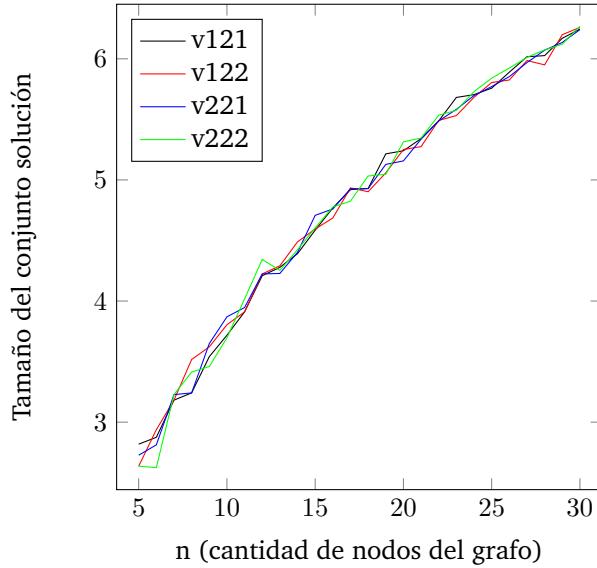
En las siguientes experimentaciones, veremos los resultados de las diferentes versiones medidos en la **cantidad de nodos** que tiene la aproximación a CIDM que devuelven los algoritmos. Por eso, los que cuenten con menor valor en el eje 'y' serán las mejores versiones:



Parecería ser que las soluciones de las versiones v121 y v122 son las mejores hasta ahora por tener menos nodos. Comparemos las demás versiones:



De estas últimas, las versiones v221 y v222 dan los mejores resultados. Veamos cuál es la diferencia que hay entre estas y las versiones v121 y v122:



Como no podemos ver grandes diferencias en los resultados obtenidos, concluimos que la mejor versión es la v221. No solamente está entre las que dieron soluciones más cercanas al CIDM de forma constante, sino que también tiene menor tiempo de ejecución.

## 7. Ejercicio 6 - Experimentación final

*Realizar una experimentación sobre un conjunto nuevo de instancias para observar la performance de los métodos comparando nuevamente la calidad de las soluciones obtenidas y los tiempos de ejecución en función de los parámetros de la entrada.*

En este ejercicio se busca comparar los tiempos de ejecución de los distintos algoritmos implementados así como la calidad de las soluciones que devuelven.

Al contrario que en las experimentaciones anteriores donde se utilizaban grafos aleatorios, en esta experimentación se busco concentrarnos en distintas familias de grafos, ya que cada una hace que ciertas implementaciones funcionen mejores que otras. Las familias de grafos elegidas fueron:

- Grafos completos ( $K_n$ ).
- Grafos de nodos aislados, es decir, el complemento de los grafos completos.
- Grafos de caminos ( $P_n$ ).
- Grafos de ciclos ( $C_n$ ).

Algunas consideraciones generales:

- Los tiempos de ejecución se midieron con la biblioteca chrono y estos fueron convertidos a nanosegundos.
- Para el algoritmo Exacto se utilizaron todas las podas planteadas en el Ejercicio 2.
- Para el algoritmo Goloso se utilizó su implementación en listas de adyacencia.
- Para el algoritmo de Búsqueda Local se utilizó como solución inicial el algoritmo Goloso y el criterio de vecindad 2, que genera soluciones vecinas a partir de quitar  $k$  vértices y agregar hasta  $k-1$  vértices.
- Para el algoritmo GRASP, se utilizó el criterio de parada 2, que sigue buscando soluciones hasta que pasan 10 iteraciones sin que cambie el cardinal de la mejor solución. Además, se utilizaron las siguientes versiones de los algoritmos:
  - Búsqueda Local utilizando el criterio de vecindad 1, que genera soluciones vecinas a partir de quitar  $k$  vértices y agregar 1 vértice.
  - Goloso utilizando listas de adyacencia.
- Para todas las familias de grafos, nos abstuvimos de comparar el algoritmo Exacto a partir de  $n \geq 14$  ya que la cantidad de tiempo que requiere a partir de ese valor es prohibitivo.

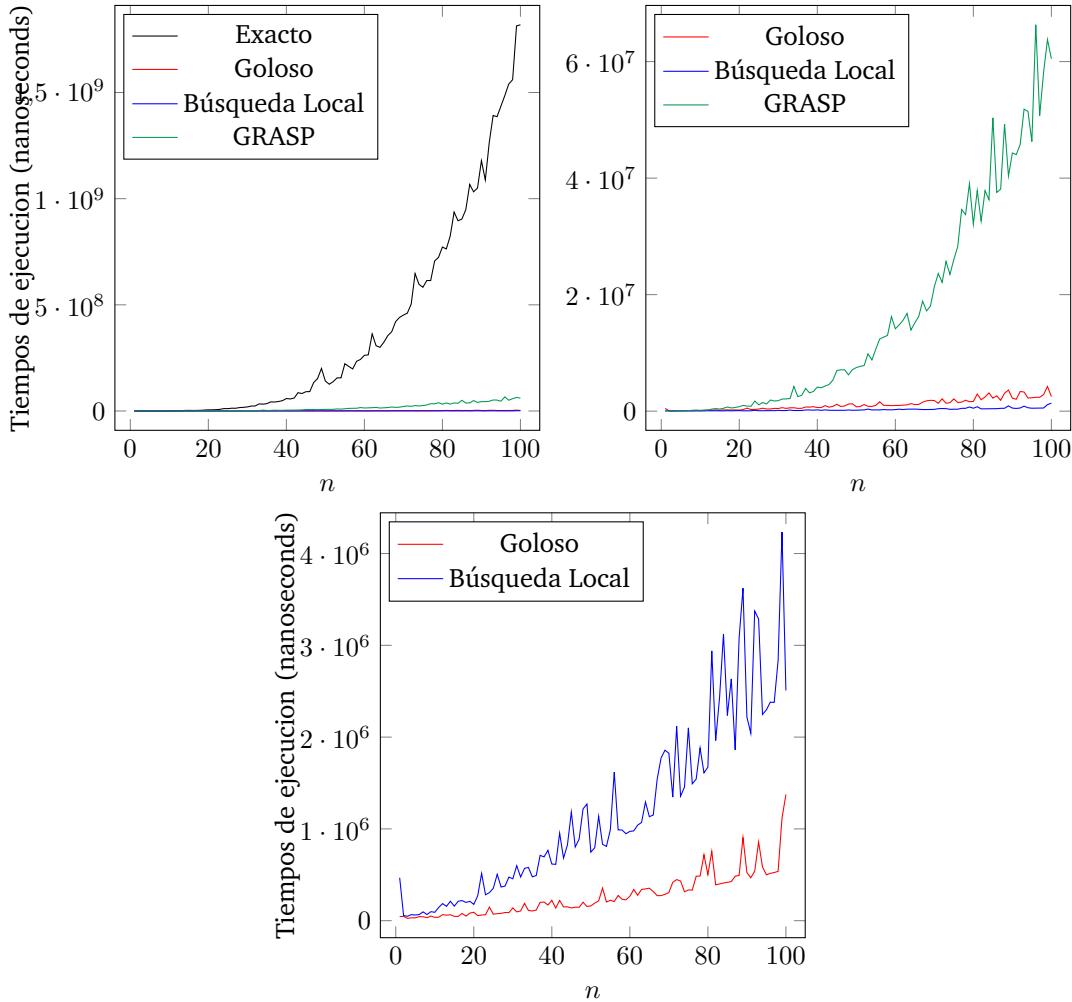
### 7.1. Experimentación con Grafos completos

Consideraciones particulares de la familia de grafos:

- Se generaron 100 casos de tests, con:
- $1 \leq n \leq 100$
- $m = \frac{n*(n-1)}{2}$ .

### 7.1.1. Tiempos de ejecución

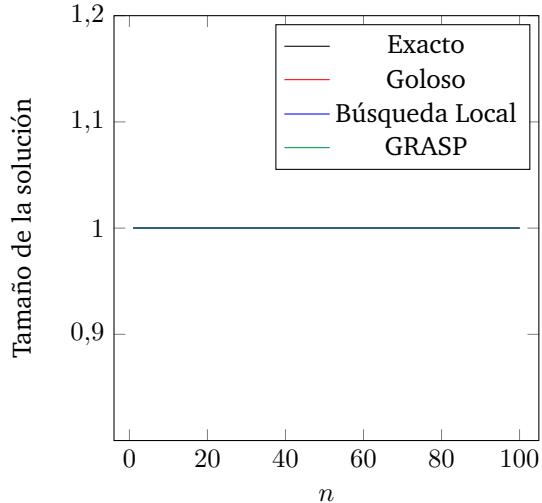
A continuación se presentan los tiempos de ejecución medidos para la familia de Grafos completos. Debido a las diferencias de magnitud entre las diferentes implementaciones, se muestran primero las 4 implementaciones juntas, luego las 3 mejores juntas y por último las 2 mejores juntas.



Como se puede ver en los gráficos presentados, para esta familia de Grafos tenemos una clara diferencia de performance:

1. En último lugar, con mayor tiempo, se encuentra el algoritmo Exacito. Esto era de esperarse ya que la complejidad exponencial de éste es mucho mayor a cualquiera de los algoritmos heurísticos. Notese aquí que pudimos experimentar con el algoritmo Exacito hasta un  $n$  grande sin mayores problemas, esto se debe a que justo esta familia de Grafos es facil de resolver para este algoritmo, no sucediendo lo mismo con las otras familias de Grafos.
2. En tercer lugar se encuentra el algoritmo GRASP, lo que era de esperarse, ya que este algoritmo utiliza en su implementación los algoritmos de Búsqueda Local y Goloso, por lo que tiene sentido que tenga un tiempo de ejecución superior a ambos. Notese aquí la diferencia de magnitud entre el Exacito y GRASP, que indica que el último es bastante más rápido ( $10^9$  para Exacito y  $10^7$  para GRASP).
3. En segundo lugar se encuentra el algoritmo de Búsqueda Local, lo que también era esperado, ya que en esta experimentación utilizamos la versión que obtiene su solución inicial utilizando un algoritmo Goloso, por lo que tiene sentido que que tenga un tiempo de ejecución superior.
4. En primer lugar, con menor tiempo, quedó el algoritmo Goloso. Esto se condice con la complejidad teórica calculada para este algoritmo, mucho menor que las demás.

### 7.1.2. Calidad de las soluciones



Para esta familia de Grafos, el tamaño de un CIDM para cualquier problema es de tamaño 1. Esto puede verse fácilmente ya que todos los nodos tienen grado  $n - 1$ , por lo que desde cualquiera de ellos se puede dominar a todos los demás, y obtenemos una solución independiente ya que no hay otros nodos en el CIDM.

Mirando el gráfico podemos ver que no hay nada que objetar en cuanto a la calidad de las soluciones obtenidas. Todos los algoritmos dieron en todos los problemas la mejor solución posible al elegir uno de los  $n$  nodos.

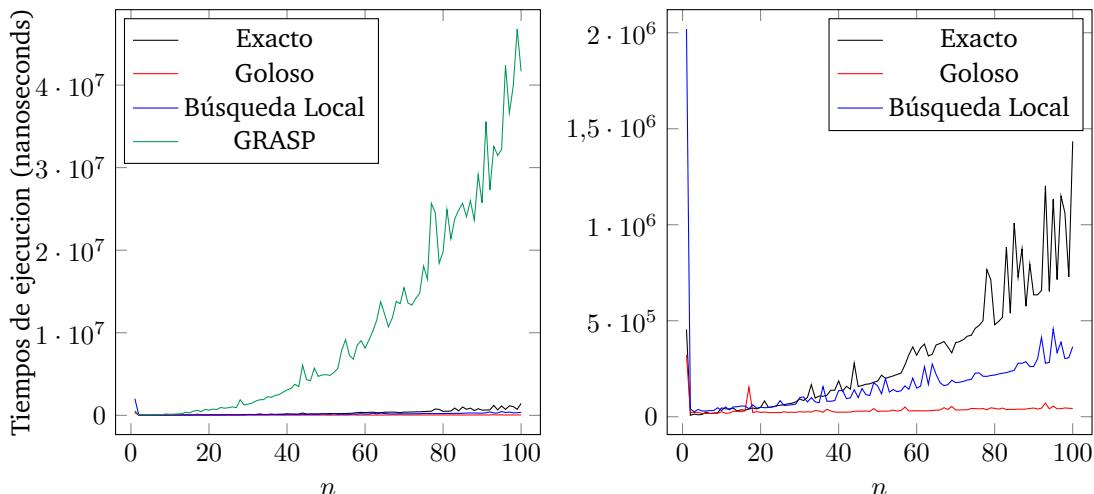
## 7.2. Experimentación con Complemento de Grafos completos

Consideraciones particulares de la familia de grafos:

- Se generaron 100 casos de tests, con:
- $1 \leq n \leq 100$
- $m = 0$ .

### 7.2.1. Tiempos de ejecución

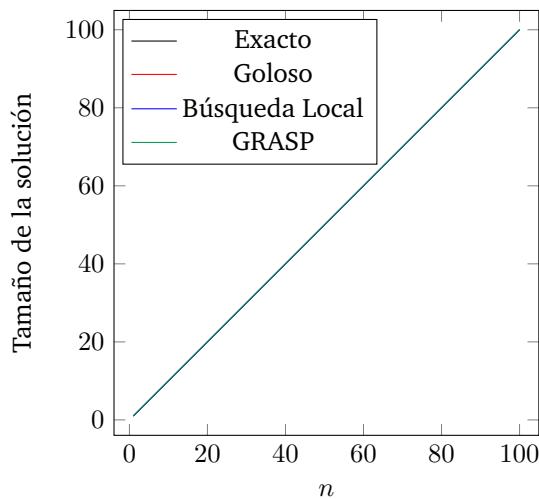
A continuación se presentan los tiempos de ejecución medidos para la familia de Grafos de nodos aislados. Debido a las diferencias de magnitud entre las diferentes implementaciones, se muestran primero las 4 implementaciones juntas y luego las 3 mejores juntas.



Como se puede ver en los gráficos presentados, para esta familia de Grafos tenemos una clara diferencia de performance. Muchos de los resultados son analogos a los de la familia anterior, por lo que se omiten. A destacar queda:

1. En último lugar, con mayor tiempo, se encuentra el algoritmo GRASP. Esto no era de esperarse, pero puede deberse a alguna facilidad del algoritmo Exacto con esta familia de Grafos en particular.
2. En tercer lugar se encuentra el algoritmo Exacto. Notese aquí que pudimos experimentar con el algoritmo Exacto hasta un  $n$  grande sin mayores problemas, esto se debe a que justo esta familia de Grafos es fácil de resolver para este algoritmo, no sucediendo lo mismo con las otras familias de Grafos.
3. La relación GRASP > Búsqueda Local > Goloso se mantiene.

### 7.2.2. Calidad de las soluciones



Para esta familia de Grafos, el tamaño de un CIDM para cualquier problema es de tamaño  $n$ . Esto puede verse fácilmente ya que todos los nodos tienen grado 0 (nodos aislados), por lo que para dominarlos a todos deben estar todos en el conjunto solución, y además al estar aislados cualquier subconjunto de nodos del grafo es independiente.

Mirando el gráfico podemos ver que no hay nada que objetar en cuanto a la calidad de las soluciones obtenidas. Todos los algoritmos dieron en todos los problemas la mejor solución posible al elegir todos los  $n$  nodos.

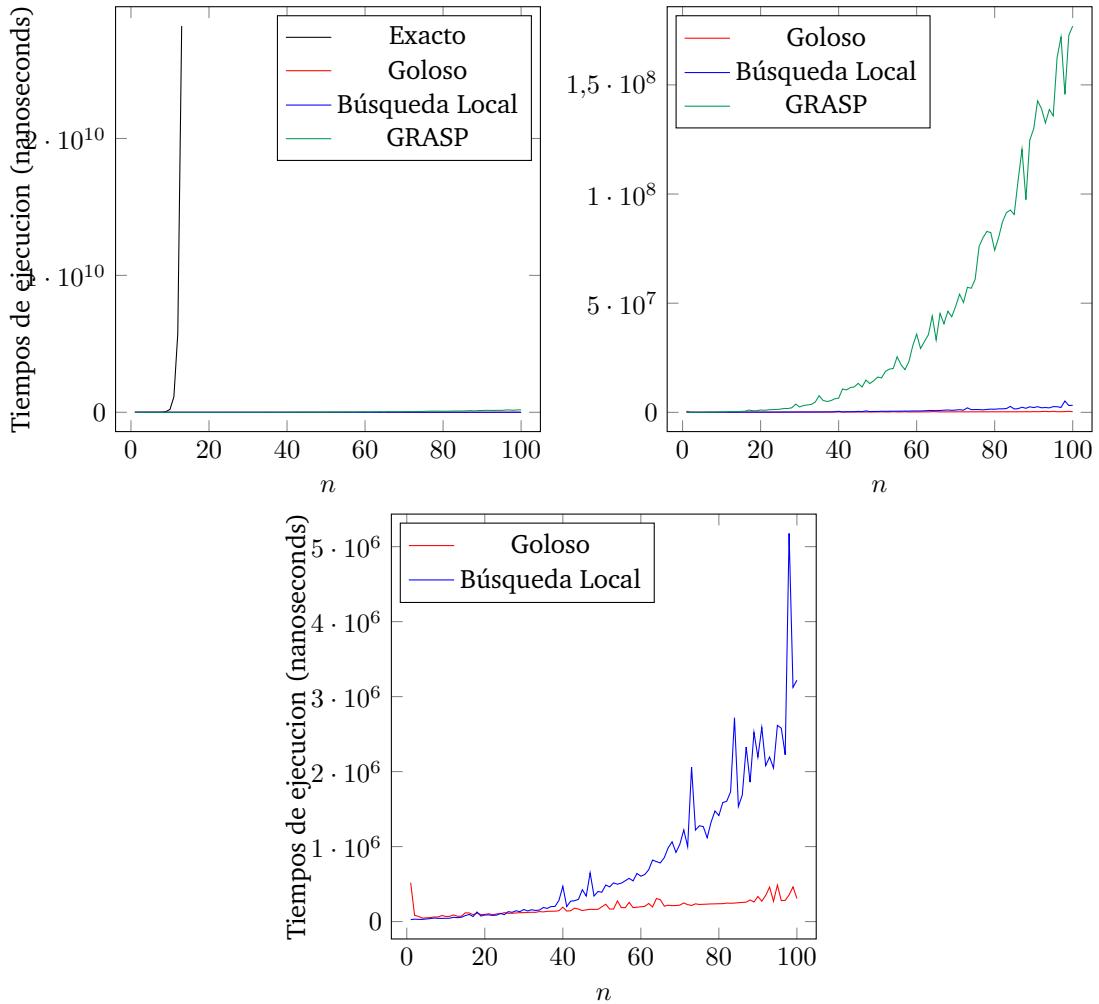
## 7.3. Experimentación con Caminos

Consideraciones particulares de la familia de grafos:

- Se generaron 100 casos de tests, con:
- $1 \leq n \leq 100$
- $m = n - 1$ .

### 7.3.1. Tiempos de ejecución

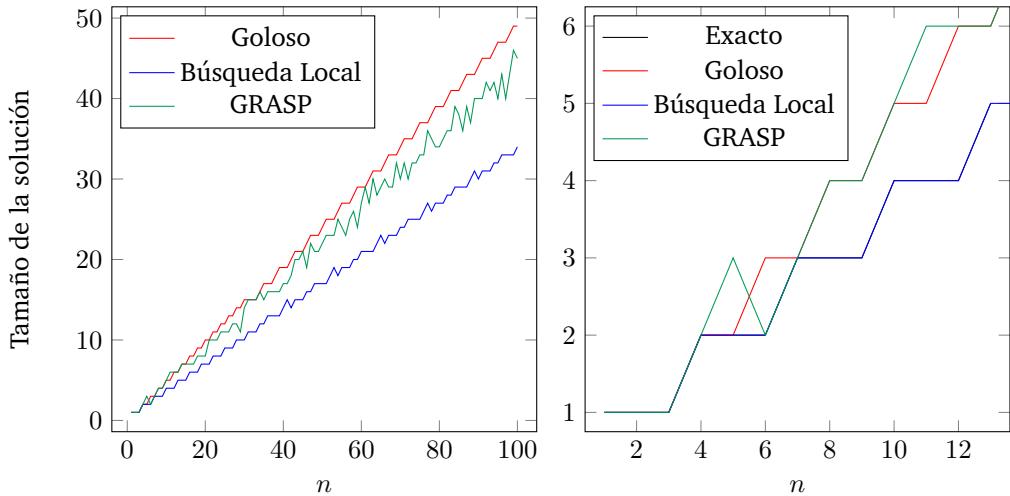
A continuación se presentan los tiempos de ejecución medidos para la familia de Grafos Camino. Debido a las diferencias de magnitud entre las diferentes implementaciones, se muestran primero las 4 implementaciones juntas, luego las 3 mejores juntas y por último las 2 mejores juntas.



Como se puede ver en los gráficos presentados, para esta familia de Grafos tenemos una clara diferencia de performance. Muchos de los resultados son analogos a los de la familia anterior, por lo que se omiten. A destacar queda:

1. En último lugar, con mayor tiempo, se encuentra el algoritmo Exacto. Notese aquí que los tiempos de este algoritmo crecen en forma desproporcionada a todos los demás, quedando evidenciada la complejidad exponencial. Para esta familia de Grafos se limitó el algoritmo Exacto a  $n = 13$ .
2. En tercer lugar, se encuentra el algoritmo GRASP. En esta familia, la diferencia entre este algoritmo y los algoritmos Goloso y Búsqueda Local es más pronunciada. Esto muy probablemente se deba al criterio de parada elegido y que la solución inicial generada por el Goloso randomizado quizás no sea tan buena. Si esto fuera cierto, entonces al algoritmo le llevaría muchas iteraciones mejorar la solución hasta que no se pueda más.
3. La relación  $\text{GRASP} > \text{Búsqueda Local} > \text{Goloso}$  se mantiene.

### 7.3.2. Calidad de las soluciones



Para esta familia de Grafos, el tamaño de un CIDM para cualquier problema es de tamaño  $\lceil n/3 \rceil$ . Esto puede verse facilmente ya que en un camino podemos elegir cada 3 nodos el del 'medio' de esos 3 para que domine a los otros 2.

Mirando el gráfico podemos ver que la mejor calidad de soluciones se obtuvo utilizando el algoritmo de Búsqueda Local (viendose además que para los primeros 13 grafos,  $0 \leq n \leq 13$ , su solución tiene el mismo tamaño que la del algoritmo Exacto) y que la peor calidad se obtuvo utilizando el algoritmo Goloso.

Al ver la calidad de las soluciones de GRASP, podemos ver que son mejores que las del algoritmo Goloso pero peores que las de Búsqueda Local. Lo primero tiene bastante sentido ya que el algoritmo GRASP parte de una solución inicial dada por un algoritmo Goloso, a la cual le aplica luego Búsqueda Local. Lo segundo no era esperado y la única explicación que encontramos es que la solución Golosa randomizada que utiliza el algoritmo GRASP como solución inicial sea consistentemente 'peor' que la solución Golosa estandar que utiliza el algoritmo de Búsqueda Local, ya que a la hora de mejorar la solución ambos utilizan el mismo criterio de búsqueda. Si esto no fuera así, el algoritmo GRASP debería poder mejorar la solución despues de varias iteraciones, pero no parece ocurrir.

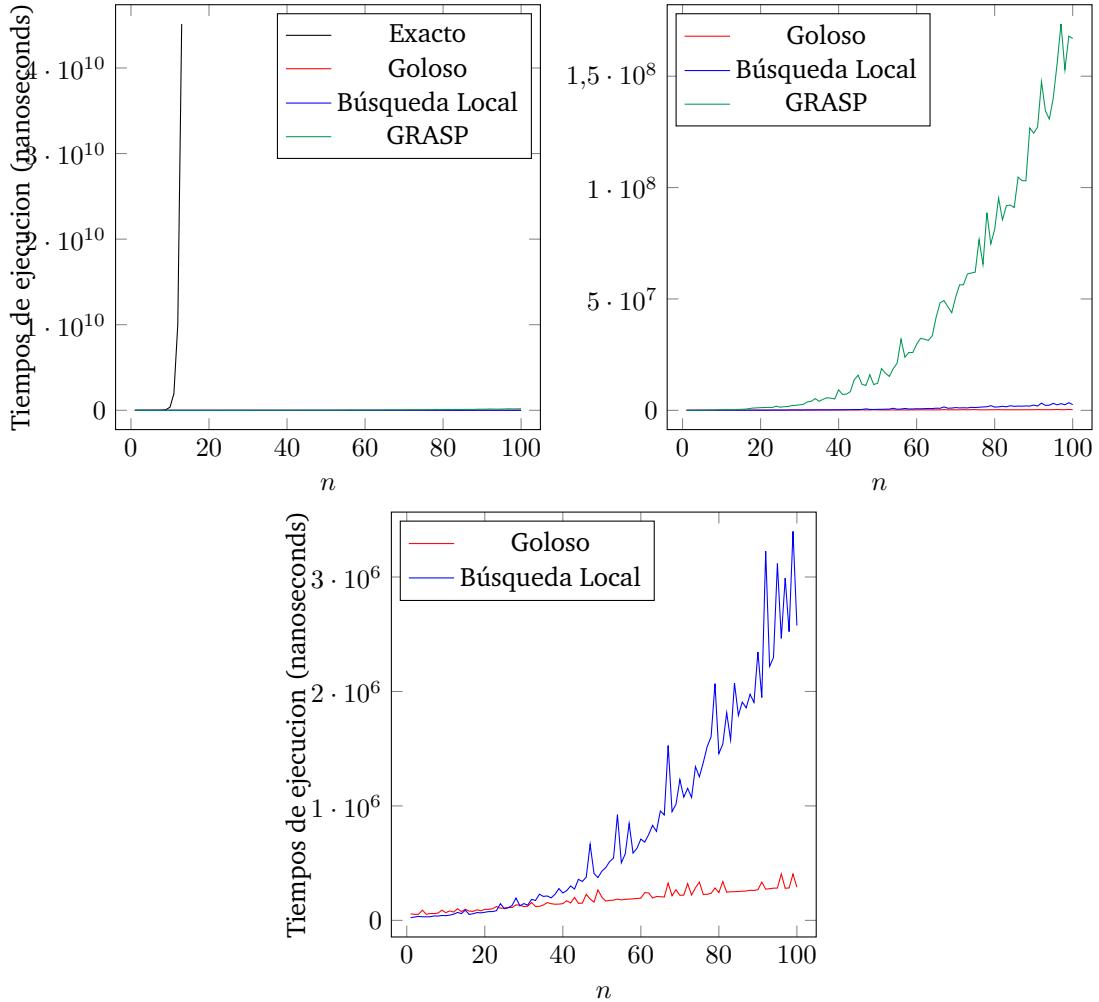
## 7.4. Experimentación con Ciclos

Consideraciones particulares de la familia de grafos:

- Se generaron 100 casos de tests, con:
- $1 \leq n \leq 100$
- $m = n$ .

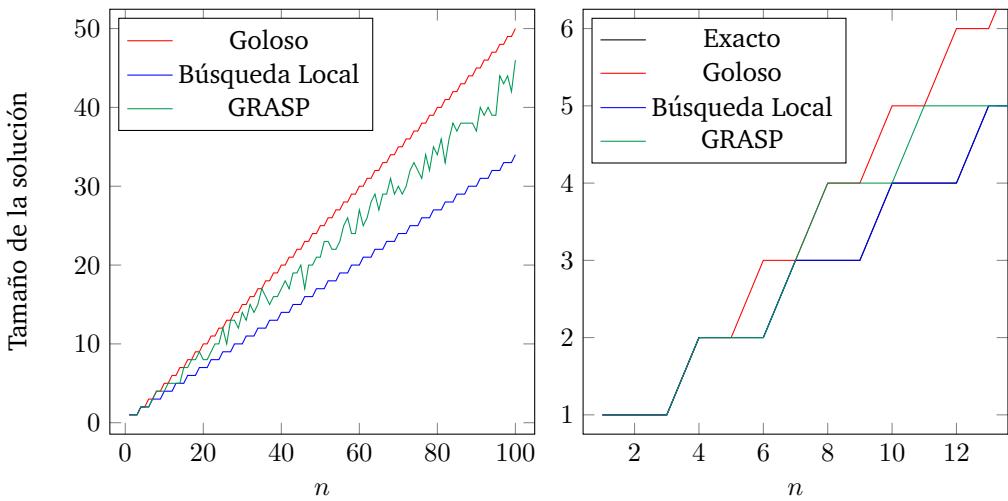
### 7.4.1. Tiempos de ejecución

A continuación se presentan los tiempos de ejecución medidos para la familia de Grafos Ciclos. Debido a las diferencias de magnitud entre las diferentes implementaciones, se muestran primero las 4 implementaciones juntas, luego las 3 mejores juntas y por último las 2 mejores juntas.



Como se puede ver en los gráficos presentados, para esta familia de Grafos tenemos una clara diferencia de performance. En esta familia se obtuvieron los mismos resultados relativos que para la familia anterior, por lo que el análisis es análogo y se omite para no repetir.

#### 7.4.2. Calidad de las soluciones



Para esta familia de Grafos, el tamaño de un CIDM para cualquier problema es de tamaño  $\lceil n/3 \rceil$ . Esto puede verse facilmente ya que en un ciclo, al igual que la familia anterior, podemos elegir cada 3 nodos el del 'medio' de esos 3 para que domine a los otros 2.

En esta familia se obtuvieron los mismos resultados relativos que para la familia anterior, por lo que el análisis es análogo y se omite para no repetir.

## 8. Conclusión

Una vez analizados los resultados en las distintas familias de grafos que fueron experimentadas, podemos llegar a las siguientes conclusiones sobre las mismas:

- El algoritmo goloso es el más rápido en tiempos de ejecución, sin embargo, la solución arrojada no siempre es la óptima. Más aún, como vimos en el ejercicio 3, la solución del goloso aunque rápida puede ser muy mala.
- El algoritmo GRASP es una buena solución media, aunque sus tiempos de ejecución son más lentos que el goloso, sus soluciones suelen ser mejores.
- Para las familias evaluadas, el mejor algoritmo fue el de búsqueda local ya que arroja no solo los mejores resultados sino unos tiempos de ejecución muy aceptables y sólo superados por los de la solución golosa, que presenta peores resultados.
- Para todo grafo con  $n \leq 13$  lo más conveniente sería usar el algoritmo exacto, ya que aunque es exponencial y mucho más lento que los otros, cuando la entrada es manejablemente pequeña uno se puede garantizar tener el resultado óptimo en un tiempo aceptable.

En conclusión, podemos ver que los algoritmos heurísticos son de gran ayuda para encarar problemas NP ya que, aunque no siempre arrojan la solución óptima, trabajan en tiempo polinomial, acortando así enormemente los tiempos de ejecución del programa. Más aún, pudimos ver que las soluciones arrojadas por algunos de los algoritmos como búsqueda local y GRASP además de correr en tiempo polinomial arrojan soluciones que pueden ser óptimas o cercanas a las óptimas. Y, aunque esto no significa tener el resultado exacto, nos permite encarar un problema exponencial y conseguir una solución rápidamente.

## A. Implementación algoritmo exacto

```
#include <vector>
#include <iostream>
#include <stack>
#include <algorithm>
using namespace std;

#define INFINITO -1

typedef vector<int> Vec;
typedef vector<Vec> Matriz;

vector<int> resolver(int n, Matriz matriz);
vector<int> resolver_aux(int n, Matriz matriz, vector<int> dom, vector<int> cidm);

// Implementacion.

int main() {
    int n, m;
    cin >> n;
    cin >> m;

    Matriz matriz(n, Vec(n, 0));

    int v, w;
    for (int i = 0; i < m; i++) {
        cin >> v;
        cin >> w;
        matriz[v-1][w-1] = 1;
        matriz[w-1][v-1] = 1;
    }

    vector<int> cidm = resolver(n, matriz);

    cout << cidm.size() << "\u25aa";
    for (int i = 0; i < cidm.size(); i++) {
        cout << cidm[i] << "\u25aa";
    }
    cout << endl;

    return 0;
}

vector<int> resolver(int n, Matriz matriz) {
    // inicializo el primer conjunto a evaluar con todos los vertices
    // del grafo
    // Este primer conjunto es dominante (pero no necesariamente
    // independiente)
    vector <int> dom(n, 0);
    for (int i = 0; i < n; i++) {
        dom[i] = i;
    }
}
```

```

        return resolver_aux(n, matriz, dom, dom);
    }

vector<int> resolver_aux(int n, Matriz matriz, vector<int> dom, vector<
    int> cidm) {
    int dom_size = dom.size();
    int cidm_size = cidm.size();
    if (cidm_size == 1) {
        return cidm;
    } else {
        // chequeo si dom es independiente
        bool indep = true;
        for (int i = 0; i < dom_size; i++) {
            for (int j = i+1; j < dom_size; j++) {
                if (matriz[dom[i]][dom[j]] == 1) {
                    indep = false;
                    break;
                }
            }
            if (!indep) {
                break;
            }
        }
    }

    if (indep) {
        // Sabemos que si Dom es dominante e independiente,
        // entonces cualquier subconjunto de Dom es no-dominante
        // Luego, devolvemos el que tiene menor cardinal entre dom y
        // cidm
        if (dom_size < cidm_size) {
            return dom;
        } else {
            return cidm;
        }
    }
}

for (int i = 0; i < dom_size; i++) {
    // copio dom y borro el i-esimo nodo de la copia
    // (no es el nodo numero i, sino el nodo en la posicion i del
    // vector)
    vector<int> copia(dom);
    copia.erase(copia.begin() + i);
    // chequeo si la copia es dominante
    bool copia_dominante = true;
    for (int i = 0; i < n; i++) {
        // chequeo si el nodo i esta en copia o es adyacente a
        // alguno en copia
        bool nodo = false;
        for (int j = 0; j < copia.size(); j++) {
            if (copia[j] == i) {
                // el nodo i esta en copia
                nodo = true;
                break;
            }
        }
        if (nodo) {
            // chequeo si el nodo i es dominante
            bool dominante = true;
            for (int k = 0; k < copia.size(); k++) {
                if (copia[k] != i && copia[k] != j) {
                    if (matriz[copia[i]][copia[k]] == 1) {
                        dominante = false;
                        break;
                    }
                }
            }
            if (!dominante) {
                copia_dominante = false;
                break;
            }
        }
    }
    if (!copia_dominante) {
        break;
    }
}
}

```

```
        } else if (matriz[i][copia[j]] == 1) {
            // el nodo i es adyacente a alguno en copia
            nodo = true;
            break;
        }
    }
    if (!nodo) {
        copia_dominante = false;
        break;
    }
}
// Sabemos que si Copia no es dominante, entonces ningun
// subconjunto de Copia es dominante,
// por lo que ni siquiera los evaluo
if (copia_dominante) {
    vector<int> nuevo_cidm = resolver_aux(n, matriz, copia, cidm
);
    if (nuevo_cidm.size() < cidm_size) {
        cidm = nuevo_cidm;
        cidm_size = nuevo_cidm.size();
    }
}
return cidm;
}
```

## B. Implementación algoritmo goloso

### B.1. Implementación sobre matriz de adyacencia

```
#include <vector>
#include <iostream>
#include <algorithm>
using namespace std;

struct Nodo {
    int numero;
    int grado;
    Nodo() : numero(0), grado(0) {};
    Nodo(const int n, const bool v, const int g) : numero(n), grado(g)
    {};
};

struct orden
{
    inline bool operator() (const Nodo& n1, const Nodo& n2)
    {
        return (n1.grado > n2.grado);
    }
};

typedef vector<int> Vec;
typedef vector<Vec> Matriz;
typedef vector<Nodo> Nodos;

vector<int> resolver(int n, Matriz matriz, Nodos nodos);

// Implementacion.
int main() {
    int n, m;
    cin >> n;
    cin >> m;

    Matriz matriz(n, Vec(n, 0));
    Nodos nodos(n, Nodo());
    for(int i = 0; i < n; i++) {
        nodos[i].numero = i;
    }

    int v, w;
    for (int i = 0; i < m; i++) {
        cin >> v;
        cin >> w;
        matriz[v-1][w-1] = 1;
        matriz[w-1][v-1] = 1;
        nodos[v-1].grado = nodos[v-1].grado + 1;
        nodos[w-1].grado = nodos[w-1].grado + 1;
    }

    vector<int> cidm = resolver(n, matriz, nodos);
}
```

```
        cout << cidm.size() << "u";
    for (int i = 0; i < cidm.size(); i++) {
        cout << cidm[i]+1 << "u";
    }
    cout << endl;

    return 0;
}

vector<int> resolver(int n, Matriz matriz, Nodos nodos) {
    // Ordeno a los nodos segun su grado
    std::sort(nodos.begin(), nodos.end(), orden());

    vector<bool> visitado(n, false);
    vector<int> cidm;
    cidm.reserve(n);

    for(int i = 0; i < n ; i++){
        if(visitado[nodos[i].numero] == false){
            visitado[nodos[i].numero] = true;
            cidm.push_back(nodos[i].numero);
            for(int j = 0; j<n; j++){
                if(matriz[nodos[i].numero][j] == 1){
                    visitado[j] = true;
                }
            }
        }
    }

    return cidm;
}
```

## B.2. Implementación sobre listas de adyacencia

```
#include <vector>
#include <iostream>
#include <algorithm>
#include <stack>
#include <list>
#include <queue>
using namespace std;

struct Nodo {
    int numero;
    int grado;
    Nodo() : numero(0), grado(0) {};
    Nodo(const int n,const bool v,const int g) : numero(n) , grado(g)
    {};
};

struct orden
{
    inline bool operator() (const Nodo& n1, const Nodo& n2)
```

```
{  
    return (n1.grado > n2.grado);  
}  
};  
  
typedef vector<Nodo> Nodos;  
typedef list<int> listaAdyacencia;  
typedef vector<listaAdyacencia> Grafo;  
  
vector<int> resolver(int n, Grafo G, Nodos nodos);  
  
// Implementacion.  
int main() {  
    int n, m;  
    cin >> n;  
    cin >> m;  
  
    Grafo G(n, listaAdyacencia());  
  
    Nodos nodos(n, Nodo());  
    for(int i = 0; i < n; i++) {  
        nodos[i].numero = i;  
    }  
  
    int v, w;  
    for (int i = 0; i < m; i++) {  
        cin >> v;  
        cin >> w;  
        G[v-1].push_back(w-1);  
        G[w-1].push_back(v-1);  
        nodos[v-1].grado = nodos[v-1].grado + 1;  
        nodos[w-1].grado = nodos[w-1].grado + 1;  
    }  
  
    vector<int> cidm = resolver(n, G, nodos);  
  
    cout << cidm.size() << "\u2014";  
    for (int i = 0; i < cidm.size(); i++) {  
        cout << cidm[i]+1 << "\u2014";  
    }  
    cout << endl;  
  
    return 0;  
}  
  
vector<int> resolver(int n, Grafo G, Nodos nodos) {  
    // Ordeno a los nodos segun su grado  
    std::sort(nodos.begin(), nodos.end(), orden());  
  
    vector<bool> visitado(n, false);  
    vector<int> cidm;  
    cidm.reserve(n);
```

```
for(int u = 0; u < n ; u++){
    if(visitado[nodos[u].numero] == false){
        visitado[nodos[u].numero] = true;
        cidm.push_back(nodos[u].numero);
        for (list<int>::iterator itAdyU=G[nodos[u].numero].begin();
            itAdyU != G[nodos[u].numero].end(); ++itAdyU) {
            int j = *itAdyU;
            cout << j << endl ;
            visitado[j] = true;
        }
    }
}

return cidm;
```

## C. Implementación algoritmo de búsqueda local

```
#include <vector>
#include <iostream>
#include <stack>
#include <list>
#include <queue>
#include <algorithm>
using namespace std;

#define VISITADO 1
#define NO_VISITADO 0
#define INCLUIDO 1
#define NO_INCLUIDO 0

// Esta versión tiene vecindades k-1 o k-k-1 vértices de diferencia
struct Nodo {
    int numero;
    int grado;
    Nodo() : numero(0), grado(0) {};
    Nodo(const int n, const int g) : numero(n), grado(g) {};
};

struct orden
{
    inline bool operator() (const Nodo& n1, const Nodo& n2)
    {
        return (n1.grado > n2.grado);
    }
};

typedef vector<Nodo> Nodos;
typedef list<int> listaAdyacencia;
typedef vector<listaAdyacencia> Grafo;

void bfs_cdi(Grafo& G, vector<int>& solucionInicial, int inicio);
void goloso(Grafo& G, vector<int>& solucionInicial);
bool solucion_posible(Grafo& G, vector<int>& solucionAuxiliar, int cantCambios);
void vecindad_primer_criterio(Grafo& G, vector<int>& solucionInicial);
void vecindad_segundo_criterio(Grafo& G, vector<int>& solucionInicial);
void imprimir_resultado(vector<int>& cidm);

// Implementacion.
int main() {
    int n, m;
    cin >> n;
    cin >> m;

    Grafo G(n, listaAdyacencia());

    int v, w;
    for (int i = 0; i < m; i++) {
```

```

        cin >> v;
        cin >> w;
        G[v-1].push_back(w-1);
        G[w-1].push_back(v-1);
    }

    vector<int> solucionInicial(n, NO_INCLUIDO);
    //Solucion Inicial 1: empiezo por el nodo 0
    //int inicio = 0;
    //bfs_cdi(G, solucionInicial, inicio);
    // Solucion Inicial 2: Heuristica golosa
    goloso(G, solucionInicial);
    //vecindad_primer_criterio(G, solucionInicial);

    vecindad_segundo_criterio(G, solucionInicial);

    imprimir_resultado(solucionInicial);

    return 0;
}

void bfs_cdi(Grafo& G, vector<int>& solucionInicial, int inicio) {
    // O(m + n)
    int n = G.size();
    vector<int> estado(n, NO_VISITADO);
    // Primera componente conexa
    estado[inicio] = VISITADO;
    solucionInicial[inicio] = INCLUIDO;
    queue<int> cola;
    cola.push(inicio);

    while(!cola.empty()) {
        int v = cola.front();
        cola.pop();
        solucionInicial[v] = INCLUIDO;
        for (list<int>::iterator it=G[v].begin(); it != G[v].end(); ++it) {
            int w = *it;
            if (solucionInicial[w] == INCLUIDO) {
                solucionInicial[v] = NO_INCLUIDO;
            }
            if (estado[w] == NO_VISITADO) {
                estado[w] = VISITADO;
                cola.push(w);
            }
        }
    }
    // Resto de las componentes conexas
    for (int u = 0; u < n; u++) {
        if (estado[u] == NO_VISITADO) {
            estado[u] = VISITADO;
            solucionInicial[u] = INCLUIDO;
            queue<int> cola;
            cola.push(u);
        }
    }
}

```

```
        while(!cola.empty()) {
            int v = cola.front();
            cola.pop();
            solucionInicial[v] = INCLUIDO;
            for (list<int>::iterator it=G[v].begin(); it != G[v].end()
                (); ++it) {
                int w = *it;
                if (solucionInicial[w] == INCLUIDO) {
                    solucionInicial[v] = NO_INCLUIDO;
                }
                if (estado[w] == NO_VISITADO) {
                    estado[w] = VISITADO;
                    cola.push(w);
                }
            }
        }
    }

void goloso(Grafo& G, vector<int>& solucionInicial) {
    int n = G.size();
    Nodos nodos(n, Nodo());

    for(int u = 0; u < n; u++) {
        nodos[u].numero = u;
        nodos[u].grado = G[u].size();
    }

    sort(nodos.begin(), nodos.end(), orden());

    vector<bool> visitado(n, false);

    for(int u = 0; u < n ; u++){
        if(visitado[nodos[u].numero] == false){
            visitado[nodos[u].numero] = true;
            solucionInicial[nodos[u].numero] = INCLUIDO;
            for (list<int>::iterator itAdyU=G[nodos[u].numero].begin();
                itAdyU != G[nodos[u].numero].end(); ++itAdyU) {
                int v = *itAdyU;
                visitado[v] = true;

            }
        }
    }
}

void vecindad_primer_criterio(Grafo& G, vector<int>& solucionInicial) {
    // Criterio de Vecindad 1: Cambiamos k vértices por 1 vértice,
    // donde k > 1
    int n = G.size();
```

```

// Genero soluciones vecinas
for (int u = 0; u < n; u++) {
    vector<int> solucionAuxiliar = solucionInicial;
    if (solucionInicial[u] == NO_INCLUIDO && G[u].size() > 1) {
        int cantINCLUIDOS = 0;
        solucionAuxiliar[u] = INCLUIDO;
        // Me fijo si el vértice NO INCLUIDO tiene al menos
        // dos vectores adyacentes INCLUIDOS
        for (list<int>::iterator itAdyU=G[u].begin(); itAdyU != G[u].end()
            ; ++itAdyU) {
            int v = *itAdyU;
            if (solucionInicial[v] == INCLUIDO) {
                cantINCLUIDOS++;
                solucionAuxiliar[v] = NO_INCLUIDO;
            }
        }
        // Necesito al menos 2 INCLUIDOS
        if (cantINCLUIDOS > 1) {
            int cantCambiosPosibles = 0;
            bool esSolucion = solucion_posible(G, solucionAuxiliar,
                cantCambiosPosibles);
            if (esSolucion) {
                // Encontre una solución Vecina mejor, fin del ciclo
                vecindad_primer_criterio(G, solucionAuxiliar);
                solucionInicial = solucionAuxiliar;
                break;
            }
        }
    }
}

void vecindad_segundo_criterio(Grafo& G, vector<int>& solucionInicial) {
    // Criterio de Vecindad 2: Cambiamos k vértices por, a lo sumo,
    // k-1 vértices, , donde k > 1
    int n = G.size();
    // Genero soluciones vecinas
    for (int u = 0; u < n; u++) {
        vector<int> solucionAuxiliar = solucionInicial;
        if (solucionInicial[u] == NO_INCLUIDO && G[u].size() > 1) {
            // Para los INCLUIDOS en la solucionInicial me fijo en sus
            // adyacentes
            // para encontrar algún adyacente que tambien esta INCLUIDO
            int cantINCLUIDOS = 0;
            solucionAuxiliar[u] = INCLUIDO;
            for (list<int>::iterator itAdyU=G[u].begin(); itAdyU != G[u].end()
                ; itAdyU++) {
                int v = *itAdyU;
                if (solucionInicial[v] == INCLUIDO) {
                    cantINCLUIDOS++;
                    solucionAuxiliar[v] = NO_INCLUIDO;
                }
            }
        }
    }
}

```

```

    }
    // Necesito al menos 2 INCLUIDOS
    if (cantINCLUIDOS > 1) {
        int cantCambiosPosibles = cantINCLUIDOS - 2;
        bool esSolucion = solucion_posible(G, solucionAuxiliar,
            cantCambiosPosibles);
        if(esSolucion){
            // Encontre una solucion Vecina mejor, fin del ciclo
            vecindad_segundo_criterio(G, solucionAuxiliar);
            solucionInicial = solucionAuxiliar;
            break;
        }
    }
}

bool solucion_posible(Grafo& G, vector<int>& solucionAuxiliar, int
cantCambios) {

    int n = G.size();
    bool esSolucion = true;
    bool finCiclo = false;

    for (int u = 0; u < n && !finCiclo; u++) {
        // Si esta INCLUIDO, sus adyacentes no pueden estar INCLUIDOS
        if (solucionAuxiliar[u] == INCLUIDO && G[u].size() > 0) {
            for (list<int>::iterator itAdyU=G[u].begin(); itAdyU != G[u].end()
                && !finCiclo; ++itAdyU) {
                int v = *itAdyU;
                if (solucionAuxiliar[v] == INCLUIDO) {
                    esSolucion = false;
                    finCiclo = true;
                }
            }
        } else if (G[u].size() > 0) {
            // Si esta NO INCLUIDO, al menos uno de sus adyacentes
            // tiene que estar INCLUIDO
            bool adyINCLUIDO = false;
            for (list<int>::iterator itAdyU=G[u].begin(); itAdyU != G[u].end()
                && !finCiclo; ++itAdyU) {
                int v = *itAdyU;
                if (solucionAuxiliar[v] == INCLUIDO) {
                    adyINCLUIDO = true;
                }
            }
            if (!adyINCLUIDO && cantCambios == 0) {
                esSolucion = false;
                finCiclo = true;
            } else if(!adyINCLUIDO) {
                // Salvo la solucion al marcar el vertice
                solucionAuxiliar[u] = INCLUIDO;
                cantCambios--;
            }
        }
    }
}

```

```
    } else {
        // Si es un K1, tiene que estar INCLUIDO
        esSolucion = solucionAuxiliar[u] == INCLUIDO;
        finCiclo = !(solucionAuxiliar[u] == INCLUIDO);
    }
}

return esSolucion;
}

void imprimir_resultado(vector<int>& cidm) {
    vector<int> solucion;
    solucion.reserve(cidm.size());
    for (int i = 0; i < cidm.size(); i++) {
        if(cidm[i] == INCLUIDO) {
            solucion.push_back(i+1);
        }
    }

    cout << solucion.size() << " ";
    for (int i = 0; i < solucion.size(); i++) {
        cout << solucion[i] << " ";
    }
    cout << endl;
}
```

## D. Implementación metaheurística GRASP

```

// INCLUDES
#include <vector>
#include <iostream>
#include <algorithm>
#include <stack>
#include <list>
#include <queue>
#include <stdlib.h>
#include <stdio.h>
using namespace std;

// DEFINES
#define VISITADO 1
#define NO_VISITADO 0
#define INCLUIDO 1
#define NO_INCLUIDO 0
#define GREEDY_RANDOM_ALPHA 0.4
#define GREEDY_RANDOM_BETA 10
#define GRASP_MAX_ITER_COUNTER 10
#define PRIMER_CRITERIO 0
#define SEGUNDO_CRITERIO 1

// ESTRUCTURAS
struct Nodo {
    int numero; //indica el indice con el que fue creado
    int grado;
    Nodo() : numero(0), grado(0) {};
    Nodo(const int n, const int g) : numero(n), grado(g) {};
};

struct orden {
    inline bool operator()(const Nodo& n1, const Nodo& n2) {
        return (n1.grado > n2.grado);
    }
};

// TYPEDEF
typedef vector<Nodo> Nodos;
typedef list<int> listaAdyacencia;
typedef vector<listaAdyacencia> Grafo;

// FUNCIONES
void recibir_parametros(Grafo& G);
void imprimir_resultado(vector<int>& cidm);

void grasp(Grafo& G, vector<int>& cidm, bool criterio_grasp, bool
criterio_greedy, bool criterio_busqueda);
void grasp_primer_criterio(Grafo& G, vector<int>& cidm, bool
criterio_greedy, bool criterio_busqueda);
void grasp_segundo_criterio(Grafo& G, vector<int>& cidm, bool
criterio_greedy, bool criterio_busqueda);

```

```
void construir_greedy_random(Grafo& G, vector<int>& solucion, bool criterio_greedy);
void construir_greedy_random_primer_criterio(Grafo& G, vector<int>& solucion);
void construir_greedy_random_segundo_criterio(Grafo& G, vector<int>& solucion);

void busqueda_local(Grafo& G, vector<int>& solucionInicial, bool criterio_busqueda);
void busqueda_local_primer_criterio(Grafo& G, vector<int>& solucionInicial);
void busqueda_local_segundo_criterio(Grafo& G, vector<int>& solucionInicial);

bool solucion_posible(Grafo& G, vector<int>& solucionCambiar, int cantCambios);
int get_indice_nodo(Nodos& nodos, int v);
int random_in_range(int min, int max);

// IMPLEMENTACION
int main() {
    srand(time(0)); // configuro el seed del random generator

    Grafo G;
    vector<int> cidm;

    recibir_parametros(G);
    grasp(G, cidm, PRIMER_CRITERIO, PRIMER_CRITERIO, PRIMER_CRITERIO);
    imprimir_resultado(cidm);

    return 0;
}

void recibir_parametros(Grafo& G) {
    int n, m;
    cin >> n;
    cin >> m;

    G = Grafo(n, listaAdyacencia());

    int v, w;
    for (int i = 0; i < m; i++) {
        cin >> v;
        cin >> w;
        G[v-1].push_back(w-1);
        G[w-1].push_back(v-1);
    }
}

void imprimir_resultado(vector<int>& cidm) {
    vector<int> solucion;
    solucion.reserve(cidm.size());
    for (int i = 0; i < cidm.size(); i++) {
```

```
        if(cidm[i] == INCLUIDO) {
            solucion.push_back(i+1);
        }
    }

    cout << solucion.size() << " ";
    for (int i = 0; i < solucion.size(); i++) {
        cout << solucion[i] << " ";
    }
    cout << endl;
}

void grasp(Grafo& G, vector<int>& cidm, bool criterio_grasp, bool criterio_greedy, bool criterio_busqueda) {
    if (criterio_grasp == PRIMER_CRITERIO) {
        grasp_primer_criterio(G, cidm, criterio_greedy,
                               criterio_busqueda);
    } else {
        grasp_segundo_criterio(G, cidm, criterio_greedy,
                               criterio_busqueda);
    }
}

void grasp_primer_criterio(Grafo& G, vector<int>& cidm, bool criterio_greedy, bool criterio_busqueda) {
    vector<int> mejor_solucion;
    construir_greedy_random(G, mejor_solucion, criterio_greedy);
    // Criterio de parada 1: hace tantas iteraciones como nodos en el grafo.
    for (int i = 0; i < G.size(); i++) {
        // Construir Solucion Greedy Random
        vector<int> nueva_solucion;
        construir_greedy_random(G, nueva_solucion, criterio_greedy);
        // Hacer busqueda local
        busqueda_local(G, nueva_solucion, criterio_busqueda);
        // Actualizar Mejor Solucion
        if (nueva_solucion.size() < mejor_solucion.size()) {
            mejor_solucion = nueva_solucion;
        }
    }
    cidm = mejor_solucion;
}

void grasp_segundo_criterio(Grafo& G, vector<int>& cidm, bool criterio_greedy, bool criterio_busqueda) {
    vector<int> mejor_solucion;
    construir_greedy_random(G, mejor_solucion, criterio_greedy);
    // Criterio de parada 2: sigue iterando hasta que no mejore la mejor solucion durante 'GRASP_MAX_ITER_COUNTER' cantidad de ciclos.
    int counter = GRASP_MAX_ITER_COUNTER;
    while (counter > 0) {
        // Construir Solucion Greedy Random
        vector<int> nueva_solucion;
        construir_greedy_random(G, nueva_solucion, criterio_greedy);
```

```
// Hacer busqueda local
busqueda_local(G, nueva_solucion, criterio_busqueda);
// Actualizar Mejor Solucion
if (nueva_solucion.size() < mejor_solucion.size()) {
    mejor_solucion = nueva_solucion;
    counter = GRASP_MAX_ITER_COUNTER;
} else {
    counter--;
}
cidm = mejor_solucion;
}

void construir_greedy_random(Grafo& G, vector<int>& solucion, bool criterio_greedy) {
    if (criterio_greedy == PRIMER_CRITERIO) {
        construir_greedy_random_primer_criterio(G, solucion);
    } else {
        construir_greedy_random_segundo_criterio(G, solucion);
    }
}

void construir_greedy_random_primer_criterio(Grafo& G, vector<int>& solucion) {
    // Criterio de Restricted Candidate List 1: los nodos que cumplan la condicion
    // de que su grado es por lo menos mejor_grado (de todos los nodos)
    * GREEDY_RANDOM_ALPHA.
    int n = G.size();
    Nodos nodos(n, Nodo());

    for(int u = 0; u < n; u++) {
        nodos[u].numero = u;
        nodos[u].grado = G[u].size();
    }

    sort(nodos.begin(), nodos.end(), orden());
    solucion = vector<int>(n, NO_INCLUIDO);

    int nodos_visitados = 0;
    while (nodos_visitados < n) {
        int mejor_grado = nodos[0].grado;
        int window_size = 0; // el maximo indice posible, la RCL va a ser nodos[0...window_size]
        for (int i = 0; i < nodos.size(); i++) {
            if (nodos[i].grado >= mejor_grado * GREEDY_RANDOM_ALPHA) {
                window_size = i;
            } else {
                break;
            }
        }
    }
}
```

```
int indice = random_in_range(0, min(window_size, (int)nodos.size() - 1));

int nodo = nodos[indice].numero;
solucion[nodo] = INCLUIDO;
nodos_visitados++;
nodos.erase(nodos.begin() + indice);

for (list<int>::iterator itAdyU=G[nodo].begin(); itAdyU != G[nodo].end(); itAdyU++) {
    int v = *itAdyU;
    int index = get_indice_nodo(nodos, v);
    if (index != -1) {
        // v no esta en el vector de nodos
        nodos.erase(nodos.begin() + index);
        nodos_visitados++;
    }
}
}

void construir_greedy_random_segundo_criterio(Grafo& G, vector<int>& solucion) {
    // Criterio de Restricted Candidate List 2: los nodos que cumplan la condicion
    // de estar entre los 'GREEDY_RANDOM_BETA' (definido antes) mejores nodos.
    int n = G.size();
    Nodos nodos(n, Nodo());

    for(int u = 0; u < n; u++) {
        nodos[u].numero = u;
        nodos[u].grado = G[u].size();
    }

    sort(nodos.begin(), nodos.end(), orden());

    solucion = vector<int>(n, NO_INCLUIDO);

    int nodos_visitados = 0;
    while (nodos_visitados < n) {
        int mejor_grado = nodos[0].grado;
        int window_size = GREEDY_RANDOM_BETA; // el maximo indice posible, la RCL va a ser nodos[0...window_size]
        int indice = random_in_range(0, min(window_size, (int)nodos.size() - 1));

        int nodo = nodos[indice].numero;
        solucion[nodo] = INCLUIDO;
        nodos_visitados++;
        nodos.erase(nodos.begin() + indice);

        for (list<int>::iterator itAdyU=G[nodo].begin(); itAdyU != G[nodo].end(); itAdyU++) {
```

```
        int v = *itAdyU;
        int index = get_indice_nodo(nodos, v);
        if (index != -1) {
            // v no esta en el vector de nodos
            nodos.erase(nodos.begin() + index);
            nodos_visitados++;
        }
    }
}

void busqueda_local(Grafo& G, vector<int>& solucionInicial, bool criterio_busqueda) {
    if (criterio_busqueda == PRIMER_CRITERIO) {
        busqueda_local_primer_criterio(G, solucionInicial);
    } else {
        busqueda_local_segundo_criterio(G, solucionInicial);
    }
}

void busqueda_local_primer_criterio(Grafo& G, vector<int>& solucionInicial) {
    // Criterio de Vecindad 1: Cambiamos k vertices por 1 vertice, donde
    // k > 1
    int n = G.size();
    // Genero soluciones vecinas
    for (int u = 0; u < n; u++) {
        vector<int> solucionAuxiliar = solucionInicial;
        if (solucionInicial[u] == NO_INCLUIDO && G[u].size() > 1) {
            int cantINCLUIDOS = 0;
            solucionAuxiliar[u] = INCLUIDO;
            // Me fijo si el vertice NO INCLUIDO tiene al menos dos vectores
            // adyacentes INCLUIDOS
            for (list<int>::iterator itAdyU=G[u].begin(); itAdyU != G[u].end()
                ; ++itAdyU) {
                int v = *itAdyU;
                if (solucionInicial[v] == INCLUIDO) {
                    cantINCLUIDOS++;
                    solucionAuxiliar[v] = NO_INCLUIDO;
                }
            }
            // Necesito al menos 2 INCLUIDOS
            if (cantINCLUIDOS > 1) {
                int cantCambiosPosibles = 0;
                bool esSolucion = solucion_posible(G, solucionAuxiliar,
                    cantCambiosPosibles);
                if (esSolucion) {
                    // Encontre una solucion Vecina mejor, fin del ciclo
                    busqueda_local_primer_criterio(G, solucionAuxiliar);
                    solucionInicial = solucionAuxiliar;
                    break;
                }
            }
        }
    }
}
```

```
        }
    }
}

void busqueda_local_segundo_criterio(Grafo& G, vector<int>&
    solucionInicial) {
    // Criterio de Vecindad 2: Cambiamos k vertices por, a lo sumo, k-1
    // vertices, , donde k > 1
    int n = G.size();
    // Genero soluciones vecinas
    for (int u = 0; u < n; u++) {
        vector<int> solucionAuxiliar = solucionInicial;
        if (solucionInicial[u] == NO_INCLUIDO && G[u].size() > 1) {
            // Para los INCLUIDOS en la solucionInicial me fijo en sus
            // adyacentes para encontrar algun adyacente que tambien esta
            // INCLUIDO
            int cantINCLUIDOS = 0;
            solucionAuxiliar[u] = INCLUIDO;
            for (list<int>::iterator itAdyU=G[u].begin(); itAdyU != G[u].end()
                ; itAdyU++) {
                int v = *itAdyU;
                if (solucionInicial[v] == INCLUIDO) {
                    cantINCLUIDOS++;
                    solucionAuxiliar[v] = NO_INCLUIDO;
                }
            }
            // Necesito al menos 2 INCLUIDOS
            if (cantINCLUIDOS > 1) {
                int cantCambiosPosibles = cantINCLUIDOS - 2;
                bool esSolucion = solucion_posible(G, solucionAuxiliar,
                    cantCambiosPosibles);
                if(esSolucion){
                    // Encontre una solucion Vecina mejor, fin del ciclo
                    busqueda_local_segundo_criterio(G, solucionAuxiliar);
                    solucionInicial = solucionAuxiliar;
                    break;
                }
            }
        }
    }
}

bool solucion_posible(Grafo& G, vector<int>& solucionCambiar, int
    cantCambios) {
    int n = G.size();
    bool esSolucion = true;
    bool finCiclo = false;
    list<int> verticesCambiados;

    for (int u = 0; u < n && !finCiclo; u++) {
        // Si esta INCLUIDO, sus adyacentes no pueden estar INCLUIDOS
        if (solucionCambiar[u] == INCLUIDO && G[u].size() > 0) {
```

```
    for (list<int>::iterator itAdyU=G[u].begin(); itAdyU != G[u].end()
        && !finCiclo; ++itAdyU) {
        int v = *itAdyU;
        if (solucionCambiar[v] == INCLUIDO) {
            esSolucion = false;
            finCiclo = true;
        }
    }
} else if (G[u].size() > 0) {
    // Si esta NO INCLUIDO, al menos uno de sus adyacentes tiene que
    // estar INCLUIDO
    bool adyINCLUIDO = false;
    for (list<int>::iterator itAdyU=G[u].begin(); itAdyU != G[u].end()
        && !finCiclo; ++itAdyU) {
        int v = *itAdyU;
        if (solucionCambiar[v] == INCLUIDO) {
            adyINCLUIDO = true;
        }
    }
    if (!adyINCLUIDO && cantCambios == 0) {
        esSolucion = false;
        finCiclo = true;
    } else if(!adyINCLUIDO) {
        // Salvo la solucion al marcar el vertice
        solucionCambiar[u] = INCLUIDO;
        cantCambios--;
    }
} else {
    // Si es un K1, tiene que estar INCLUIDO
    esSolucion = solucionCambiar[u] == INCLUIDO;
    finCiclo = !(solucionCambiar[u] == INCLUIDO);
}

return esSolucion;
}

int get_indice_nodo(Nodos& nodos, int v) {
    int n = nodos.size();
    for (int i = 0; i < n; i++) {
        if (nodos[i].numero == v) {
            return i;
        }
    }
    return -1;
}

int random_in_range(int min, int max) {
    return min + (rand() % (max - min + 1));
}
```