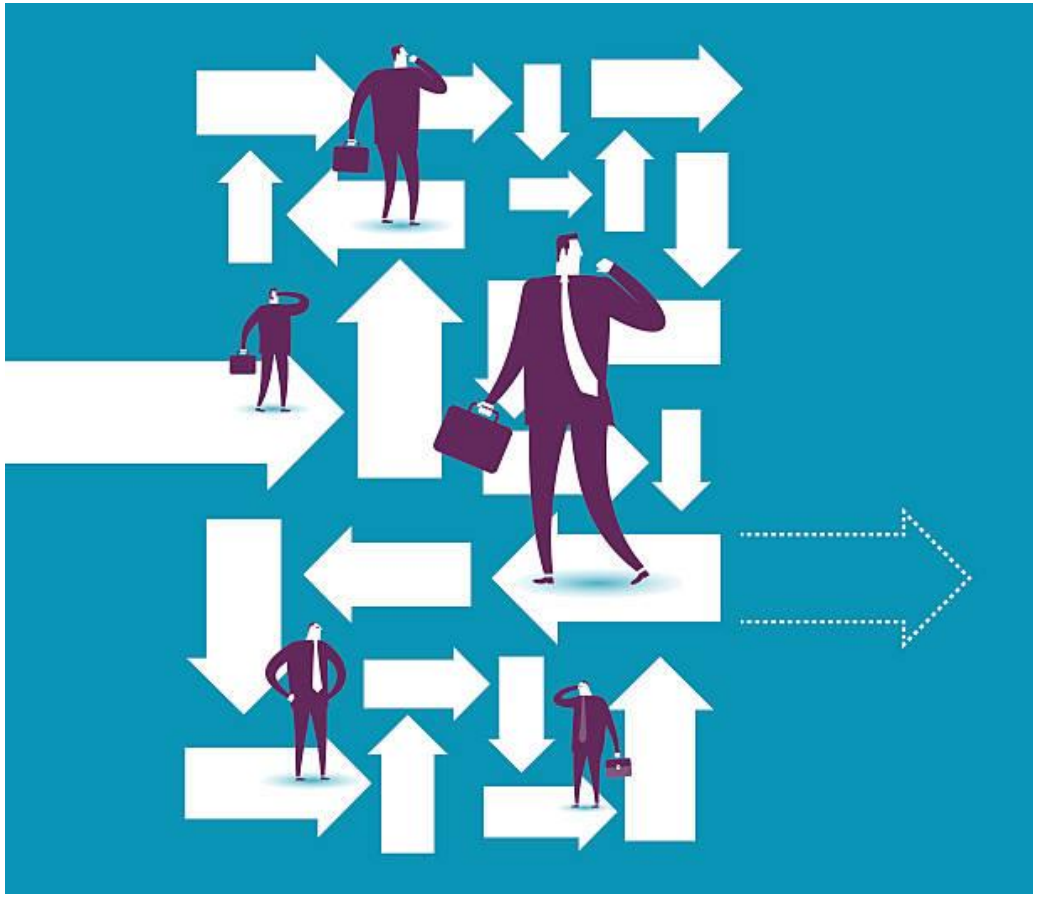


Práctica 4

Exploración de Grafos

MARIO LÍNDEZ MARTÍNEZ

JUAN AYUSO ARROYAVE

MARIO MARTÍN RODRÍGUEZ

INDICE

- Objetivos.
- Definición del problema.
- Algoritmo Diseñado.
- Eficiencia.
- Análisis comparativo de rendimiento y tiempo para distintas funciones de cota.
- Análisis comparativo de las dos técnicas heurísticas.
- Conclusión.

Objetivos

El objetivo de esta práctica es el uso y compresión de técnicas para resolver problemas basados en la exploración de grafos, concretamente las técnicas “Backtracking” y “Branch and Bound”. Para ello, hemos resuelto el problema del viajante de comercio utilizando cada una de estas dos técnicas que hemos nombrado.

Definición del problema

Enunciado

Un agente comercial tiene que visitar n ciudades. Se proporciona una matriz cuadrada D, de dimensiones n × n tal que di,j indica el coste de viajar desde la ciudad i a la j. El agente debe realizar un recorrido visitando todas las ciudades y finalizando en la misma ciudad desde la que partió. Además, sólo puede visitar cada ciudad una sola vez. El coste global de un recorrido dado es la suma de los costes de todos los traslados. El objetivo es encontrar un recorrido que cumpla las restricciones establecidas y tenga el menor coste posible.

Algoritmo Diseñado. Backtracking

Realización

Nosotros hemos trabajado en la implementación del algoritmo con la clase BT (Backtracking), a la cual le hemos añadido diversos métodos que nos ayudarán a resolver este problema.

En primer lugar, cargamos con el método load la matriz de distancias, el vector que contendrá el índice de salidas mínimas de cada fila (usado para el cálculo de la Cota2), la COTA_GLOBAL (menor distancia a través de un método Greedy) e inicializamos la solución a la obtenida por el método Greedy (heurística usada la del vecino más cercano).

Código del método load:

```
1. ifstream file;
2.         file.open(filename);
3.
4.         if (!file){
5.             cout << "ERROR ABRIENDO EL ARCHIVO" << endl;
6.             exit(-1);
7.         }
8.
9.         int nodos_aux;
10.        file >> nodos_aux;
11.
12.        NUM_NODOS = nodos_aux;
13.
14.        reservaMemoria(NUM_NODOS);
15.
16.        for (int i = 0; i < NUM_NODOS; ++i) {
17.            for (int j = 0; j < NUM_NODOS; ++j) {
18.                int num;
19.                file >> num;
20.
21.                distancias[i][j] = num;
22.            }
23.        }
24.
25.        file.close();
30.
31.        vector<bool> ya_pertenece (NUM_NODOS);
32.        ya_pertenece.at(0) = true;
```

```
34.         CalculaSalidasMinimas();
35.
36.         minimo_coste = distancias[0][salidas_minimas.at(0)];
37.         for (int i = 1; i < NUM_NODOS; ++i){
38.             if (minimo_coste > distancias[i][salidas_minimas.at(i)]){
39.                 minimo_coste = distancias[i][salidas_minimas.at(i)];
40.             }
41.         }
42.
43.         pair<vector<int>, int> sol_greedy = Greedy (distancias, ya_pertenece, NUM_NODOS);
44.         COTA_GLOBAL = sol_greedy.second;
45.         solucion = sol_greedy.first;
46.
```

Luego de esto, pasamos al método más importante, el cual se trata de pvc. Este método nos servirá como acceso a la resolución del problema usando las diferentes cotas.

El funcionamiento tanto de pvc1 como de pvc2 es equivalente, sin embargo, cada uno usará la primera cota local o la segunda respectivamente. Esto se podría haber generalizado usando punteros a funciones, pero por falta de tiempo hemos realizado una implementación más rápida, aunque para nada recomendable.

Usaremos pvc1 para explicar el funcionamiento del programa, equivalente en ambos casos. Hemos optado por un diseño recursivo, por lo que nuestra función recibirá como argumentos el nodo actual en el que se encuentra (índice fila), un vector de solución auxiliar y el coste del camino recorrido actual. En primer lugar, quitaremos el nodo actual del set que almacena los nodos que todavía quedan sin visitar y lo añadiremos al vector solución auxiliar. Tras esto comprobaremos si ya se han visitado todos los nodos, es decir, el nodo en el que estamos es una hoja de nuestro árbol de estados.

En caso de que así sea sumaremos al coste total la distancia de volver al nodo inicial (0) desde el último nodo considerado y comprobaremos si es un menor coste que lo ya almacenado. Si es menor, la consideraremos como la nueva mejor solución, tomando como nueva COTA_GLOBAL su coste.

En caso contrario, comprobaremos la cota local del nodo actual, la cual nos dará una estimación de cuánto nos costará en términos mínimos seguir por ese camino. Si estimamos que el camino merecerá la pena (que la cota local sea menor que la global), pasaremos a procesar los caminos restantes. Antes de llamar recursivamente a la función sumaremos el coste de ir a ese nuevo nodo, se procesará recursivamente y al volver deberemos retroceder al estado anterior. Esto es restar su coste considerado antes de entrar a la función, añadirlo de nuevo al set de nodos sin_visitar y eliminarlo de la solucion_auxiliar.

```
1. void pvc1 (int k, vector<int> & solucion_aux, int coste=0){
2.         sin_visitar.erase(k);
3.         solucion_aux.push_back(k);
4.
5.         if (sin_visitar.empty()){
6.             coste += distancias[k][0];
7.             if (coste < COTA_GLOBAL){
8.                 COTA_GLOBAL = coste;
9.                 solucion = solucion_aux;
10.                PintaSolucion(coste);
11.            }
12.        } else {
13.            int cota_local = CotaLocal1(coste);
14.            if (cota_local < COTA_GLOBAL){
15.                set<int> sin_visitar_aux = sin_visitar;
16.
17.                for (auto it = sin_visitar_aux.begin(); it != sin_visitar_aux.end(); ++it){
18.                    coste += distancias[k][*it];
19.                    pvc2(*it, solucion_aux, coste);
20.
21.                    coste -= distancias[k][*it];
22.                    sin_visitar.insert(*it);
23.                    solucion_aux.pop_back();
24.                }
25.            }
26.        }
27.    }
28.
```

```
1. for (int i = 0; i < NUM_PETICIONES; ++i){
2.     if (!Pertenece(inmediato, peticiones.at(i))){ // Pertenece = O(N)
3.         if (inmediato.size() < TAMANIO_INMEDIATO){
4.             inmediato.push_back(peticiones.at(i));
5.
6.         } else {
7.             inmediato.at(CancionMasLejana(inmediato, peticiones, i)) = peticiones.at(i);
8.             ++NUM_INTERCAMBIOS;
9.         }
10.
11.     }
12. }
13.
```

Justificación de la validez del algoritmo

Nuestro algoritmo es prácticamente un algoritmo de fuerza bruta optimizado ya que irá recorriendo la totalidad del árbol de estados dejándose por el camino aquellas ramas que ya estimamos que no nos ofrecerán una solución válida. Es por ello por lo que este algoritmo será valido y ofrecerá una solución correcta en cada ejecución.

Funciones de cota alternativas

Para la resolución de nuestro problema, hemos usado dos funciones de cota. La primera de ellas tomaremos la distancia más baja de toda la matriz de distancias, es decir, la salida con el coste mínimo. Dicha distancia será usada para estimar el mínimo coste de ir a los nodos que aún queden sin visitar y se sumará al coste del camino actual.
Código CotaLocal1:

```
1. int CotaLocal1 (int coste){
2.         int num_nodos_sin_visitar = sin_visitar.size();
3.
4.         coste += minimo_coste*num_nodos_sin_visitar;
5.
6.         return (coste);
7.     }
```

Funciones de cota alternativas

La segunda cota local usará el mismo concepto, pero en vez de tomar una única salida mínima tomaremos el coste mínimo especifico de salir de cada nodo. Esto se almacenó en el vector salidas_minimas.

Código CotaLocal2:

```
1. int CotaLocal2 (int coste){
2.         coste += GetSalidaMinima(solucion.back());
3.         for (auto it = sin_visitar.begin(); it != sin_visitar.end(); ++it){
4.             coste += GetSalidaMinima(*it);
5.         }
6.
7.         return (coste);
8.     }
9. int GetSalidaMinima (int k){
10.         return (distancias[k][salidas_minimas.at(k)]);
11. }
```

Eficiencia Teórica

La eficiencia del algoritmo es indeterminada, pues número de llamadas recursivas dependerán del caso en el que nos encontremos en ese momento. Por lo tanto, no tiene sentido hablar de eficiencia teórica.

Con respecto a las cotas, claramente la cota1 es O(1) ya que son dos operaciones sencillas e inmediatas, mientras que la cota2 dependerá del número de nodos que queden sin visitar, esto es O(k).

Eficiencia Empírica

Los datos que vamos a calcular a continuación son utilizando la segunda cota que hemos definido en el apartado anterior por ser la más precisa y la supuestamente más eficiente.

Para realizar un análisis de eficiencia empírica deberemos ejecutar el mismo algoritmo para diferentes tamaños de entrada. Para este algoritmo, lo ejecutaremos para diferentes tamaños del vector del número de nodos que hay. Estos tiempos los almacenamos en un fichero dist_bt.dat. El código es el siguiente:

```
#!/bin/bash
#echo "" >> salida.dat
printf "" > dist_bt.dat

i=1
while [ "$i" -le 15 ]
do
      # Generamos los puntos
      ./generador $i data_bt.txt
      printf "PUNTOS GENERADOS\t"

      # Ejecutamos los puntos
      ./bt data_bt.txt >> dist_bt.dat

      echo "Terminado $i"

      i=$(( $i + 1 ))
done
```

Eficiencia Empírica

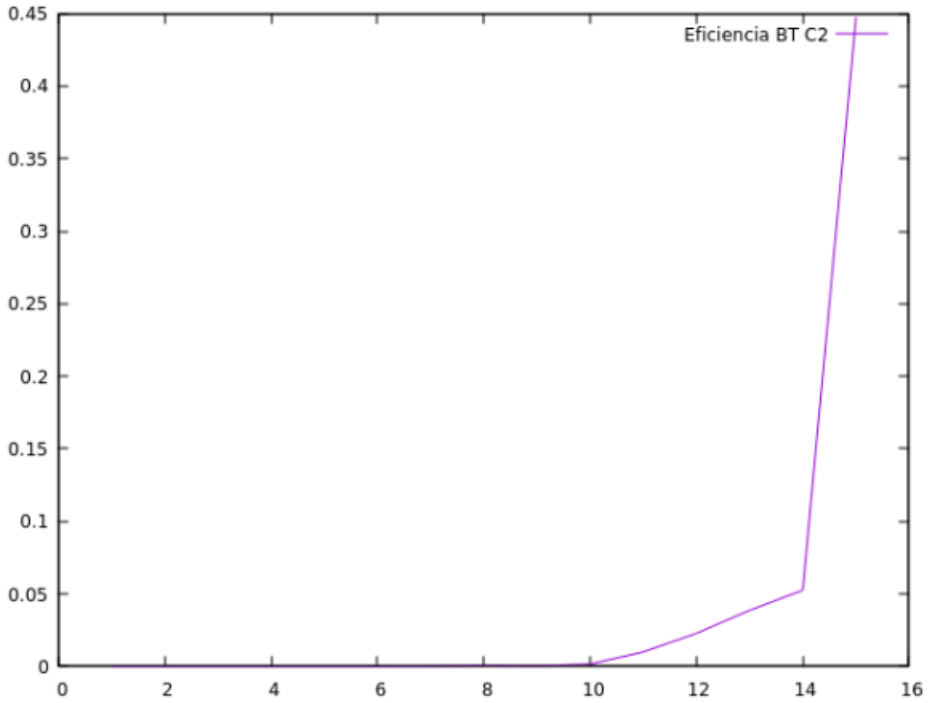
Empezamos con un tamaño base de 1 punto y vamos aumentándolo de 1 en 1 hasta llegar al tamaño de 15 puntos.

Los tiempos obtenidos son los siguientes:

Tamaño	Tiempo (seg)
1	0,00000209
2	0,00000274
3	0,00000135
4	0,00001924
5	0,00001313
6	0,00001370
7	0,00015042
8	0,00080108
9	0,00037696
10	0,00165159
11	0,01007440
12	0,02290940
13	0,03869480
14	0,05255540
15	0,44691400

Eficiencia Empírica

A partir de dichos tiempos, queda la siguiente gráfica:

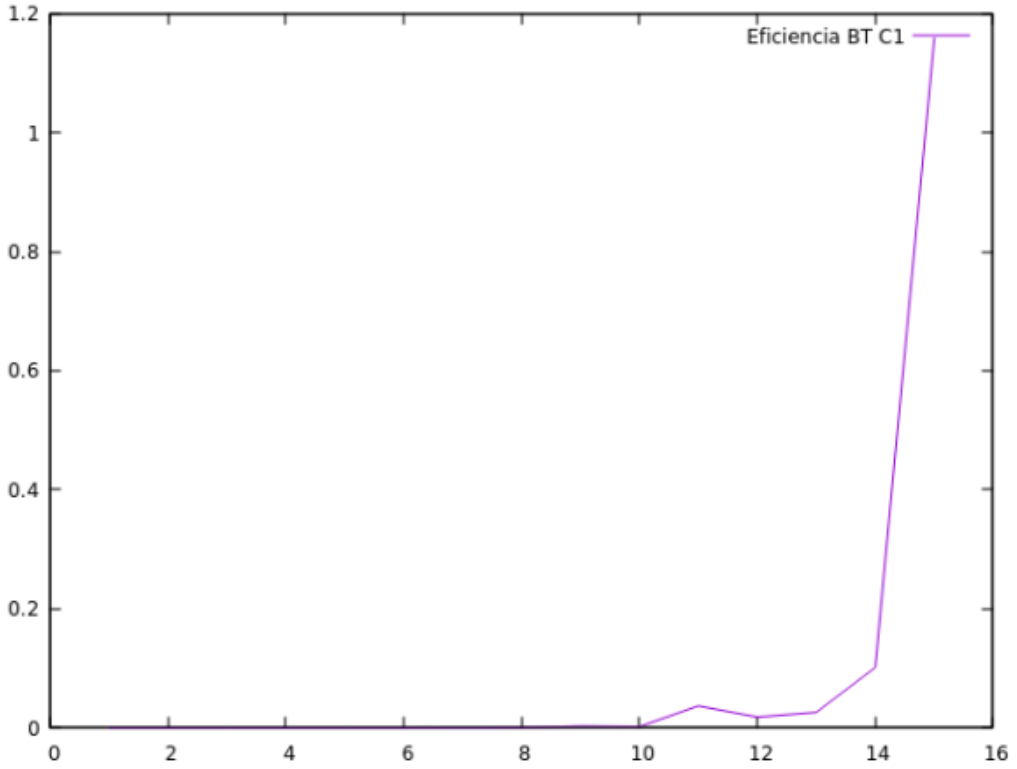


Gráfica de tiempos del algoritmo diseñado

Análisis Comparativo de Rendimiento

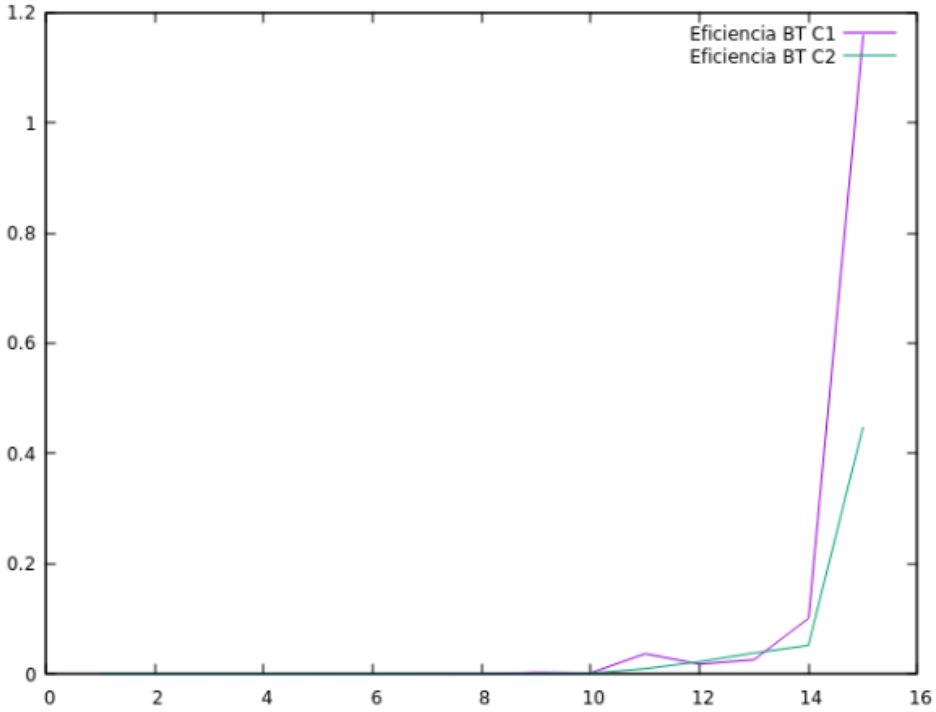
Para la otra función de cota que hemos implementado, los datos y gráfica son las siguientes:

Tamaño	Tiempo (seg)
1	0,00000095
2	0,00000267
3	0,00000449
4	0,00000400
5	0,00001897
6	0,00001128
7	0,00014896
8	0,00067671
9	0,00322674
10	0,00251811
11	0,03738360
12	0,01861960
13	0,02640090
14	0,10151000
15	1,15773000



Gráfica de tiempos del algoritmo diseñado

Si juntamos las dos gráficas de las distintas cotas, nos queda lo siguiente:



Gráfica de tiempos comparada

Donde claramente podemos observar como la ganancia en precisión de la cota2 es más determinante que su propia eficiencia, pues hace el programa más eficiente.

Algoritmo Diseñado. Branch and Bound

Realización

Para la técnica Branch and Bound nos volveremos a apoyar en una clase BB cuya implementación será muy similar a la clase BT.

Además de esto nos apoyaremos del functor Comparison para ordenar la cola con prioridad perteneciente a la clas BB, el cual comparará por el segundo elemento del pair de elementos. Éste será la cota local del nodo añadido como primer elemento.

```
1. struct Comparison {
2.     bool operator() (const pair<nodo,int>& n1, const pair<nodo,int> & n2){
3.         return (n1.second > n2.second);
4.     }
5. };
```

Realización

Continuemos viendo el resto del funcionamiento de la clase.

En primer lugar, la inicialización vendrá dada de nuevo por un método load el cual será análogo en comportamiento con el de la clase BT. Se encargará de inicializar la matriz de distancias, el nodo origen, la solución, la cota global inicial (calculadas de nuevo con el algoritmo Greedy del vecino más cercano), el vector de salidas mínimas y el mínimo coste (objetos usados para la cota2 y 1 respectivamente).

Código del método load:

Realización

```
1.    void load (const char filename[]){
2.        ifstream file;
3.        file.open(filename);
4.
5.        if (!file){
6.            cout << "ERROR ABRIENDO EL ARCHIVO" << endl;
7.            exit(-1);
8.        }
9.
10.       int nodos_aux;
11.       file >> nodos_aux;
12.
13.       NUM_NODOS = nodos_aux;
14.
15.       reservaMemoria(NUM_NODOS);
16.
17.       set<int> set_origen;
18.
19.       for (int i = 0; i < NUM_NODOS; ++i) {
20.           for (int j = 0; j < NUM_NODOS; ++j) {
21.               int num;
22.               file >> num;
23.
24.               distancias[i][j] = num;
25.           }
26.
27.           if (i != 0){
28.               set_origen.insert(i);
29.           }
30.
31.       }
32.
33.       file.close();
```

Realización

```
35.         origen = nodo(vector<int>{0}, set_origen, 0);
36.
37.         // Siempre partimos de la ciudad 0
38.         //solucion.push_back(0);
39.         //sin_visitar.erase(0);
40.
41.         vector<bool> ya_pertenece (NUM_NODOS);
42.         ya_pertenece.at(0) = true;
43.
44.         CalculaSalidasMinimas();
45.
46.         minimo_coste = distancias[0][salidas_minimas.at(0)];
47.         for (int i = 1; i < NUM_NODOS; ++i){
48.             if (minimo_coste > distancias[i][salidas_minimas.at(i)]){
49.                 minimo_coste = distancias[i][salidas_minimas.at(i)];
50.             }
51.         }
52.
53.
54.         pair<vector<int>, int> sol_greedy = Greedy (distancias, ya_pertenece, NUM_NODOS);
55.         COTA_GLOBAL = sol_greedy.second;
56.         solucion = sol_greedy.first;
57.
58.     }
```

Realización

Tras esto, veamos el funcionamiento del método pvc(), de nuevo puerta de acceso para la resolución del problema usando las distintas cotas diseñadas:

```
1. void pvc (int cota){
2.     if (cota == 1){
3.         pvc1();
4.     } else if (cota == 2){
5.         pvc2();
6.     }
7. }
```

Realización

De nuevo, esta no es una implementación recomendable, pero sí es funcional. Analicemos ahora pvc1, el cual será análogo en comportamiento a pvc2.

Para empezar, añadiremos nuestro primer nodo y punto de partida a la priority_queue caminos, la cual irá conteniendo los posibles nodos a desarrollar. Tras esto comenzaremos el algoritmo en sí el cual consistirá en ir desarrollando el mejor camino posible (información proporcionada por la priority_queue) hasta que no queden nodos prometedores (mientras caminos no este vacío).

Dentro del bucle tomaremos el tope de la cola, eliminándolo de la misma en el proceso, como el nodo más prometedor y procederemos a desarrollarlo. Si este nodo fuera una hoja, esto es que no le queden nodos por visitar) actualizaremos el coste del camino con la vuelta al nodo 0 y comprobaremos si es una mejor solución que la ya poseíamos. Si es así, la tomaremos como nueva mejor solución, tomando como cota global su coste.

Realización

En caso de que todavía le queden nodos por visitar, vemos si la cota local de dicho nodo es menor que la global, en caso de que se haya modificado esta última. Si lo es, procederemos a añadir como a la cola aquellos caminos que nos deriven del nodo tomado. Para ello, nos apoyaremos en un nodo auxiliar inicialmente igual que prometedor y le actualizaremos el nuevo camino tomado hacia cada uno de los nodos que le queden por visitar. Así mismo, añadimos también al coste la distancia desde el último nodo de la solución hasta el que acabamos de añadir. Si la cota de este nuevo posible camino es menor que la cota global, lo tomaremos en cuenta y lo añadiremos a caminos.

Código pvc1:

```
1.    void pvc1(){
2.         caminos.push(pair<nodo,int>{this->origen, CotaLocal2(this->origen)});
3.
4.         while (!caminos.empty()){
5.             nodo prometedor = caminos.top().first;
6.             caminos.pop();
7.
8.             if (prometedor.sin_visitar.empty()){
9.                 prometedor.coste += distancias[prometedor.solucion.back()][0];
10.
11.                if (prometedor.coste < COTA_GLOBAL){
12.                    solucion = prometedor.solucion;
13.                    COTA_GLOBAL = prometedor.coste;
14.                }
15.
16.            } else {
17.                if (prometedor.coste < COTA_GLOBAL){
18.                    for (auto it = prometedor.sin_visitar.begin(); it !=
prometedor.sin_visitar.end(); ++it){
19.                        nodo aux = prometedor;
20.
21.                        aux.sin_visitar.erase(*it);
22.                        aux.solucion.push_back(*it);
23.                        aux.coste += distancias[prometedor.solucion.back()][*it];
24.
25.                        int cota = CotaLocal1(aux);
26.
27.                        if (cota < COTA_GLOBAL){
28.                            caminos.push(pair<nodo,int>{aux,cota});
29.                        }
30.                    }
```

Justificación de la validez del algoritmo

Por motivos análogos al algoritmo Backtracking, este algoritmo será válido, pues irá recorriendo el árbol a través de aquellos nodos que sean útiles, encontrando así todas las mejores soluciones hasta el momento. A diferencia del algoritmo antes mencionado, Branch and Bound desarrollará antes los nodos antes de tomarlos como procesables.

Funciones de cota alternativas

Para la resolución de nuestro problema, hemos usado dos funciones de cota. Para la primera de ellas tomaremos la distancia más baja de toda la matriz de distancias, es decir, la salida con el coste mínimo. Dicha distancia será usada para estimar el mínimo coste de ir a los nodos que aún queden sin visitar y se sumará al coste del camino actual.

Código CotaLocal1:

```
1.    int CotaLocal1(const nodo & n) const{
2.         int coste = n.coste;
3.         int num_nodos_sin_visitar = n.sin_visitar.size();
4.
5.         coste += (minimo_coste*num_nodos_sin_visitar);
6.
7.         return (coste);
8.     }
9.
```

Funciones de cota alternativas

La segunda cota local usará el mismo concepto, pero en vez de tomar una única salida mínima tomaremos el coste mínimo especifico de salir de cada nodo. Esto se almacenó en el vector salidas_minimas.

Código CotaLocal2:

```
1.    int CotaLocal2 (const nodo & n){
2.         int coste = n.coste;
3.
4.         coste += GetSalidaMinima(solucion.back());
5.
6.         for (auto it = n.sin_visitar.begin(); it != n.sin_visitar.end(); ++it){
7.             coste += GetSalidaMinima(*it);
8.         }
9.
10.        return (coste);
11.    }
12.
```

Eficiencia Teórica

De nuevo, de forma análoga a la técnica Backtracking, no tiene sentido hablar de eficiencia teórica pues el número de nodos a desarrollar y procesar dependerá de cada caso.

Como las cotas usadas son realmente las mismas, mantendrán sus respectivas eficiencias, esto es $Cota1 \in O(1)$ y $Cota2 \in O(k)$, siendo k el número de nodos que le queden sin visitar al nodo actual.

Eficiencia Empírica

Para realizar un análisis de eficiencia empírica deberemos ejecutar el mismo algoritmo para diferentes tamaños de entrada. Para este algoritmo, lo ejecutaremos para diferentes tamaños del vector del número de nodos que hay. Estos tiempos los almacenamos en un fichero dist_bb.dat. El código es el siguiente:

```
1. #!/bin/bash
2. pinté "" > dist_bb.dat
3.
4. i=1
5. while [ "$i" -le 15]
6. do
7. # Generamos los puntos
8. ./generador $i data_bb.txt
9.
10. # Ejecutamos los puntos
11. ./bb data_bb.txt >> dist_bb.dat
12.
13. echo "Terminado $i"
14.
15. i=$(( $i + 1))
16. done
```

Eficiencia Empírica

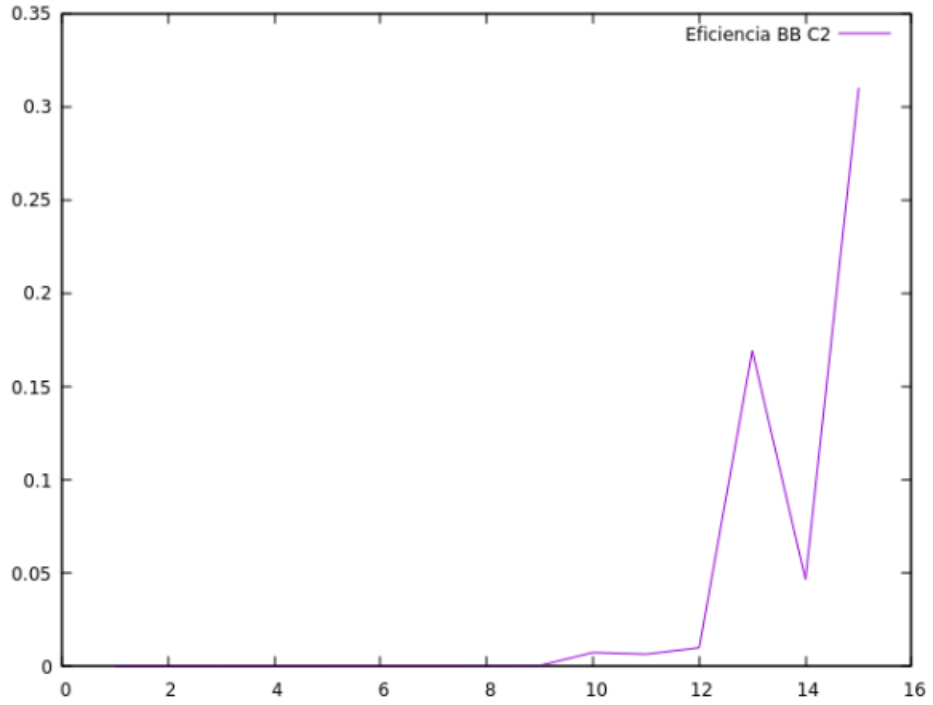
Empezamos con un tamaño base de 1 punto y vamos aumentándolo de 1 en 1 hasta llegar al tamaño de 15 puntos.

Los tiempos obtenidos son los siguientes:

Tamaño	Tiempo (seg)
1	0,00000096
2	0,00000137
3	0,00000875
4	0,00000143
5	0,00005219
6	0,00016528
7	0,00013243
8	0,00041251
9	0,00054091
10	0,00758994
11	0,00664845
12	0,01012400
13	0,16925900
14	0,04675850
15	0,30960900

Eficiencia Empírica

A partir de dichos tiempos, queda la siguiente gráfica:



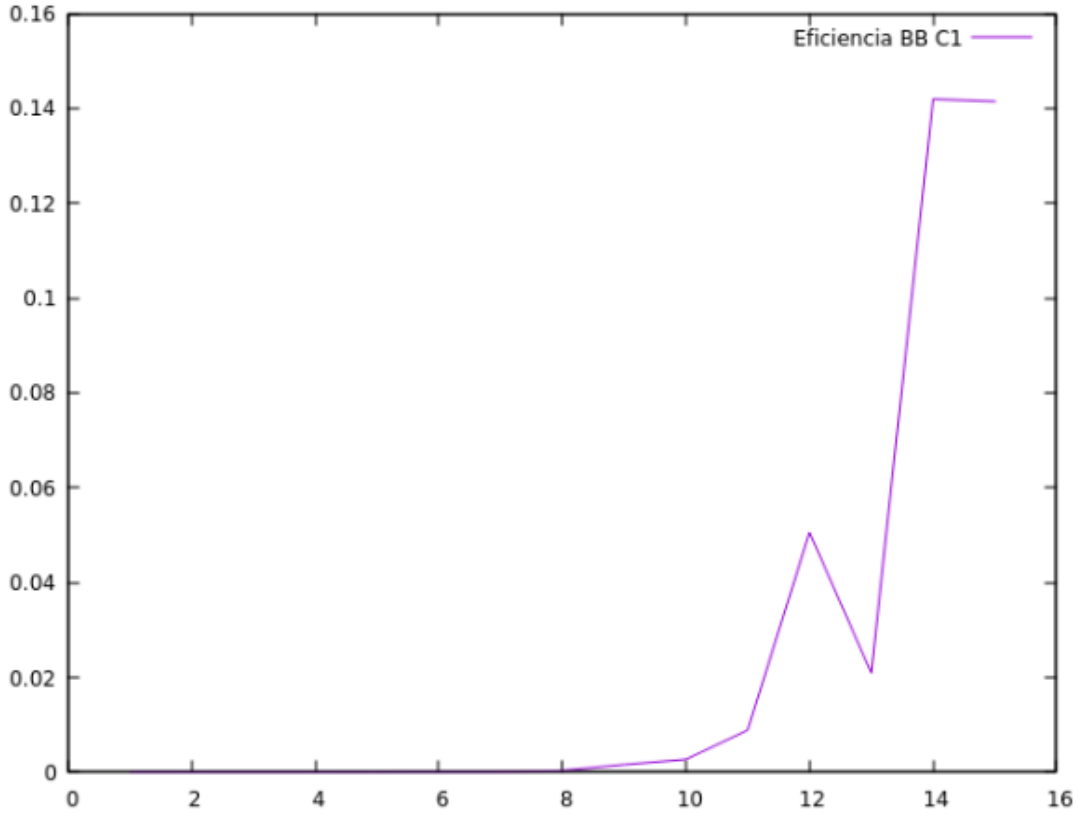
Gráfica de tiempos del algoritmo diseñado

Análisis comparativo del rendimiento

Para la otra función de cota que hemos i ...s datos y gráfica son las siguientes:

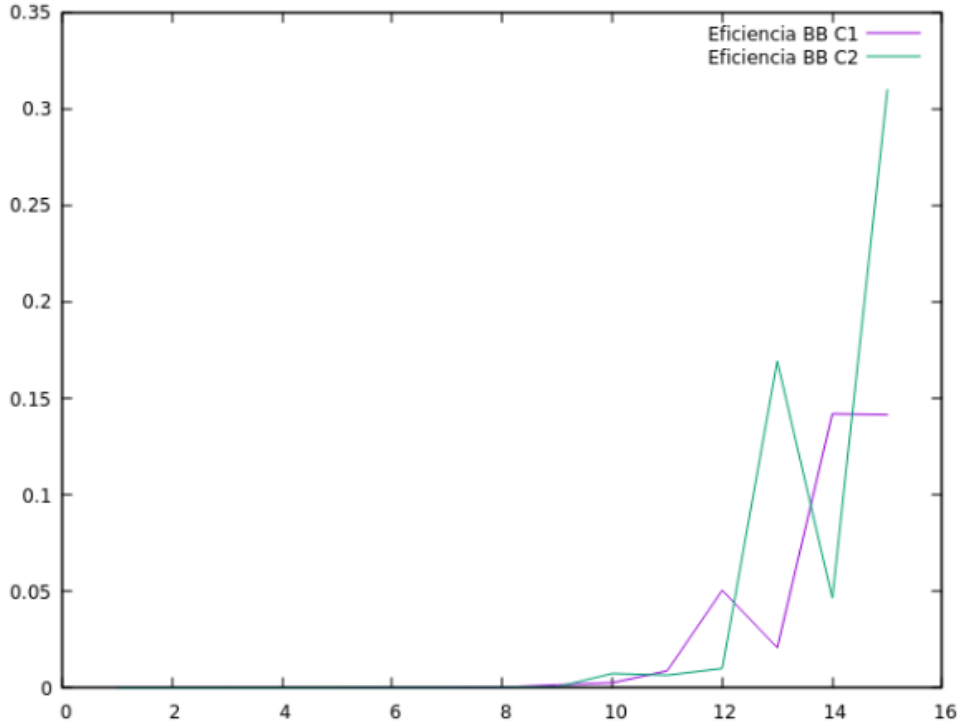
Tamaño	Tiempo (seg)
1	0,00000092
2	0,00000266
3	0,00000415
4	0,00000833
5	0,00002599
6	0,00007267
7	0,00008538
8	0,00037099
9	0,00163019
10	0,00272298
11	0,00890477
12	0,05060160
13	0,02095930
14	0,14192600
15	0,14147100

Análisis comparativo del rendimiento



Análisis comparativo del rendimiento

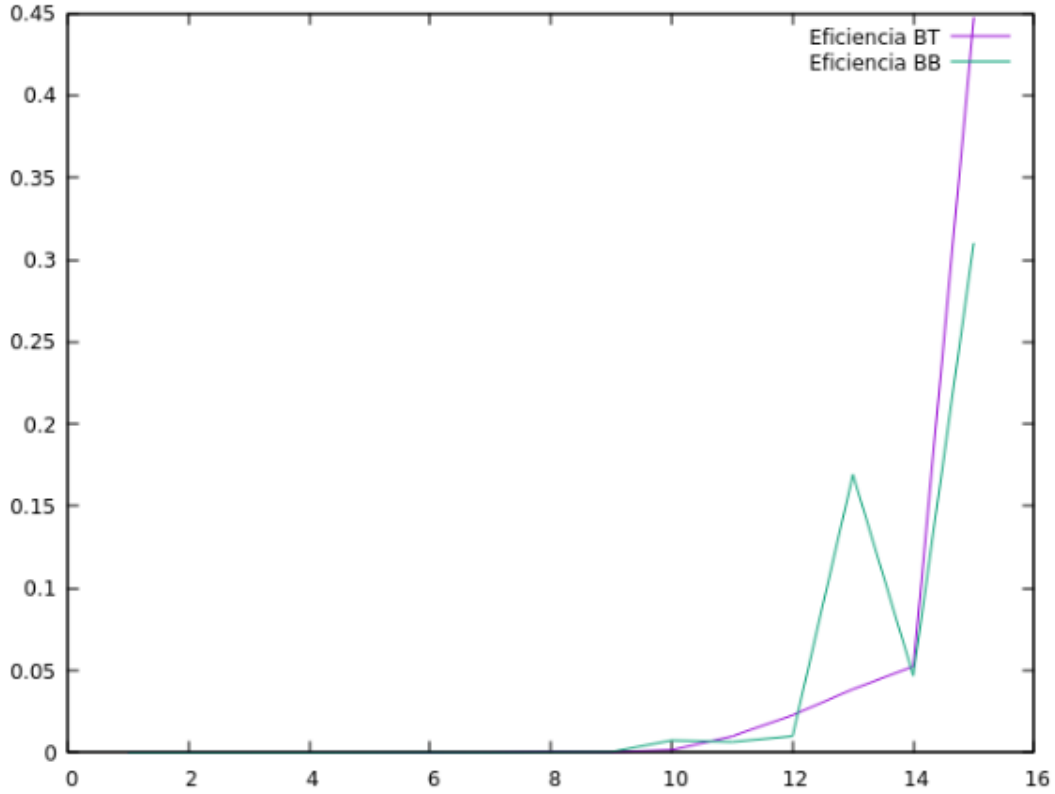
Si juntamos las dos gráficas de las distintas cotas, nos queda lo siguiente:



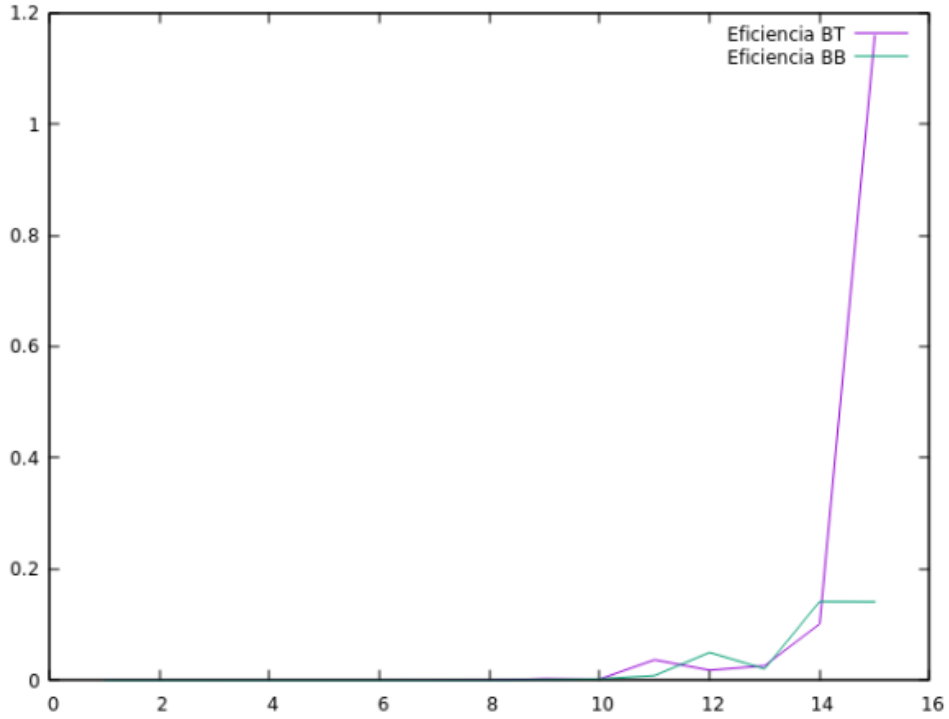
Gráfica de tiempos comparada

Comparación de las dos heurísticas

Utilizando en los dos algoritmos la primera cota, la gráfica de tiempos resultante es la siguiente:



Utilizando en los dos algoritmos la segunda cota, la gráfica de tiempos resultante es la siguiente:



Gráfica de tiempos comparada

Conclusión

Como primera conclusión, hemos visto que el diseño de la cota es un factor muy importante ya que la eficiencia de nuestro algoritmo puede mejorar dependiendo de esta.

Además, hemos visto que la técnica “Branch and Bound” es más eficiente que la técnica “Backtracking”. Esta conclusión realmente es obvia, ya que con “Backtracking” se va a comprobar todo el árbol entero de posibilidades mientras que con “Branch and Bound” vamos descartando ramas cuya información ya nos dice que no nos va a servir.