DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Práctico II

Algoritmos y Estructuras de Datos III
Primer Cuatrimestre de 2015

Integrante	LU	Correo electrónico
Aldasoro Agustina	86/13	agusaldasoro@gmail.com
Noriega Francisco	660/12	frannoriega.92@gmail.com
Zimenspitz Ezequiel	155/13	ezeqzim@gmail.com
Zuker Brian	441/13	brianzuker@gmail.com

Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Ciudad Universitaria - (Pabellón I/Planta Baja)
Intendente Güiraldes 2160 - C1428EGA
Ciudad Autónoma de Buenos Aires - Rep. Argentina
Tel/Fax: (54 11) 4576-3359
http://www.fcen.uba.ar

Resumen

Poner resumen.

Índice

1. Dakkar

1.1. Descripción de la problemática

La problemática trata de una travesía, la cual cuenta con *n* cantidad de etapas. Para cada una de las etapas, se puede elegir recorrerla en alguno de los tres vehículos disponibles: una BMX, una motocross o un buggy arenero. Cada uno de ellos permite concretar cada etapa en cantidades de tiempo diferentes. Además, la cantidad de veces que se pueden usar la motocross y el buggy arenero está acotada por k_m y k_b respectivamente.

Los *tiempos* que le llevan a los vehículos recorrer el trayecto varían por cada etapa y son datos conocidos pasados por parámetro.

Se pide recorrer la travesía, dentro de las restricciones, de modo que se utilice la menor cantidad de tiempo posible. Si existen dos (o más) maneras de atravesarla dentro del tiempo óptimo, se pide devolver sólo una.

Se exige resolver la problemática con una complejidad temporal de $O(n.k_m.k_b)$.

Dibujitos con ejemplos :)

1.2. Resolución propuesta y justificación

Para resolver esta problemática, optamos por implementar un algoritmo de *Programación Dinámica*.

Con el fin de encontrar el recorrido factible que emplee menos tiempo; debemos comparar, para cada etapa, cuál es el menor tiempo con el que puede recorrer el camino faltante eligiendo en la instancia actual uno de los tres vehículos disponibles. Dado que la formulación de este problema es muy extensa, se realizó una formulación recursiva de modo que para cada problema se le asigna un valor dependiendo de un subproblema menor.

Formulación Recursiva

Optamos por comenzar recorriendo desde la etapa n hasta la etapa 0; *n* va a indicar la etapa actual, k_m la cantidad de motos y k_b la cantidad de boogys restantes que se pueden utilizar.

- Cuando llegamos a la etapa $n = 0$ es porque terminamos todo el recorrido, de modo que el tiempo devuelto va a ser 0.
- Cuando $k_m = 0$ y $k_b = 0$ es porque la etapa actual (*n*) y el recorrido restante (las *n-1* etapas) lo vamos a tener que hacer sólo en bicicleta, sin importar el tiempo que conlleve ya que nos quedamos sin motos y boogys para usar.
- Cuando $k_m = 0$ y $k_b \neq 0$ es porque utilizamos la mayor cantidad de motos posibles y las *n-1* etapas restantes -conjunto a la actual(*n*)- las vamos a tener que recorrer con Bicicleta o Boogy. Por este motivo se elige la opción con tiempo menor usando Bicicleta o Boogy en la etapa *n* y llamando recursivamente a la función para *n-1* considerando esta elección.
- De modo análogo, cuando $k_m \neq 0$ y $k_b = 0$ sólo vamos a contar con Motos y Bicicletas para la etapa actual y las *n-1* etapas faltantes.
- En cambio, en caso contrario, todavía tenemos disponible cantidad de los tres vehículos. Por este motivo, se comparan los tres casos: empleando la Bicicleta en la etapa n, la Moto o el Boogy llamando recursivamente a la función para *n-1* de modo que va a devolver el menor tiempo posible considerando la elección llevada a cabo.

$$func(n, k_m, k_b) = \begin{cases} 0 & \text{si } n = 0 \\ tiempoBici(n) + f(n-1, 0, 0) & \text{si } k_m = 0 \wedge k_b = 0 \\ min \begin{pmatrix} tiempoBici(n) + func(n-1, 0, k_b), \\ tiempoBoogy(n) + func(n-1, 0, k_b - 1) \end{pmatrix} & \text{si } k_m = 0 \wedge k_b \neq 0 \\ min \begin{pmatrix} tiempoBici(n) + func(n-1, k_m, 0), \\ tiempoMoto(n) + func(n-1, k_m - 1, 0) \end{pmatrix} & \text{si } k_m \neq 0 \wedge k_b = 0 \\ min \begin{pmatrix} tiempoBici(n) + func(n-1, k_m, k_b), \\ tiempoMoto(n) + func(n-1, k_m - 1, k_b), \\ tiempoBoogy(n) + func(n-1, k_m, k_b - 1) \end{pmatrix} & \text{sino} \end{cases}$$

Dado que los n, k_m y k_b iniciales van a ser los dados por parámetro y en el planteo de nuestra ecuación en la llamada recursiva n siempre decrementa en 1 y los demás o bien quedan iguales o uno de ellos decrementa en uno, estos parámetros van a estar acotados por:

$$\begin{array}{rcccl} 0 & \leq & n & \leq & n_{parametro} \\ 0 & \leq & k_m & \leq & k_{m_{parametro}} \\ 0 & \leq & k_b & \leq & k_{b_{parametro}} \end{array}$$

1.3. Análisis de la complejidad

1.3.1. Complejidad Temporal

1.3.2. Complejidad Espacial

Si bien, ya no piden ningun requisito, pongamos cuanta memoria usa :)

1.4. Código fuente

1.5. Experimentación

1.5.1. Constrastación Empírica de la complejidad

2. Zombieland II

2.1. Descripción de la problemática

2.2. Resolución propuesta y justificación

2.3. Análisis de la complejidad

2.4. Código fuente

2.5. Experimentación

2.5.1. Constrastación Empírica de la complejidad

3. Refinando petróleo

3.1. Descripción de la problemática

3.2. Resolución propuesta y justificación

3.3. Análisis de la complejidad

El algoritmo recibe por parámetro un grafo no dirigido en forma de cantidad de nodos (numerados arbitrariamente de 0 a n-1) y los ejes que existen entre cada par de nodos, y el costo de construir una refinería. Es importante destacar que la cantidad de ejes está acotada por $O(n^2)$, porque cada nodo, en el peor caso, puede conectarse con todos los nodos del grafo, excepto consigo mismo, es decir que para cada nodo, existen a lo sumo n-1 ejes que entran y salen del mismo. Tomamos como notación $n = cantNodos$ y $m = cantEjes$.

Lo primero que hace es identificar los ejes donde es conveniente colocar una tubería para conectar dos pozos en vez de construir una refinería sobre ambos, aplicando el algoritmo de $Kruskal$, apoyados sobre la estructura $Union - Find$.

Este paso ordena los ejes mediante sort Poner referencia a sort si todavia no se uso de la librería standar de C. Esto tarda $O(m.log(m))$, lo que es lo mismo que $O(n^2.log(n^2))$, que por propiedades del logaritmo es $O(n^2,2.log(n))$, eliminando constantes es $O(n^2.log(n))$. Luego por cada eje de ejes, si no forma un ciclo dentro de la componente conexa a la que pertenece, une ambos pozos, pero solo lo agrega al resultado si cuesta menos que poner una refinería. Este paso aprovecha la estructura de conjuntos disjuntos $Union - Find$, para ver si dos conjuntos son dijuntos y eventualmente unirlos rápidamente". Dado que lo implementamos con las heurísticas de $pathcompression$ y $unionbyrank$, la complejidad de m operaciones find_set y union-set más una llamada a make_set de complejidad $O(n)$, ejecutan en tiempo $O(m\alpha(n))$

3.4. Código fuente

```
class UnionFind {
public:
     UnionFind(int tamano);
    ~UnionFind();
    int find_set(int x);
    void union_set(int x, int y);
    bool is_in(int x, int y);

private:
    vector<int> parent;
    vector<int> rank;
};
```

```
 UnionFind::UnionFind(int tamano){
     parent = vector<int>(tamano);
    rank = vector<int>(tamano);
//cada indice es su propio representante, su ranking es 0
    for (int i = 0; i < tamano; ++i) {
        parent[i] = i;
        rank[i] = 0;
    }
}
```

```
int UnionFind::find_set(int x) {
//si es representante lo devuelvo, si no lo hago apuntar directamente al representante
//para que llamadas consecutivas cuesten tiempo constante
    if(parent[x] != x)
        parent[x] = find_set(parent[x]);
    return parent[x];
}
```

```
void UnionFind::union_set(int x, int y) {
//buscamos representantes de ambos elementos
//requiere que no esten en el mismo conjunto
    int rx = find_set(x);
    int ry = find_set(y);
//incluyo el de menor ranking en el de mayor para mantener balanceada la estructura
    if(rank[rx] < rank[ry]){
        parent[rx] = ry;
    }
    else{
        parent[ry] = rx;
        if(rank[ry] == rank[rx])
            rank[rx]++;
    }
}
```

```
bool UnionFind::is_in(int x, int y) {
    return find_set(x) == find_set(y);
}
```

```
struct eje {
    unsigned int pozoA;
    unsigned int pozoB;
    unsigned int costoTuberia;
    bool operator< (const eje& otro) const{
        return costoTuberia < otro.costoTuberia;
    }
};
```

```
int main(int argc, char const *argv[]){
    unsigned int pozos, cantConexiones, costoRefineria;
    unsigned int pozoA, pozoB, costoTuberia;
    cin >> pozos >> cantConexiones >> costoRefineria;
    vector<eje> ejes;
//leemos la entrada, la almacenamos en ejes
    for (int i = 0; i < cantConexiones; ++i){
        cin >> pozoA >> pozoB >> costoTuberia;
        pozoA--;
        pozoB--;
        eje conex;
        conex.pozoA = pozoA;
        conex.pozoB = pozoB;
        conex.costoTuberia = costoTuberia;
        ejes.push_back(conex);
    }
//aplicamos el algoritmo
    refinandoPetroleo(ejes, pozos, costoRefineria);
    return 0;
}
```

```
int refinandoPetroleo(vector<eje>& ejes, int cantPozos, int costoRefineria){
//generamos los arboles minimos para cada componente conexa, en realidad solo los ejes que cuesten menos que
    vector<eje> conexionesMinimas = generarArbolesMinimos(ejes, costoRefineria, cantPozos);
    UnionFind conexos(cantPozos);
    int costoTotal = 0, cantRef = 0;
//armamos un union-find para identificar componentes triviales, en las demas solo hara falta poner una refin
//en el representante de la componente pues los tubos son mas baratos para unir los distintos pozos
    for (int i = 0; i < conexionesMinimas.size(); ++i) {
        conexos.union_set(conexionesMinimas[i].pozoA, conexionesMinimas[i].pozoB);
        costoTotal += conexionesMinimas[i].costoTuberia;
    }
//en las componentes triviales van refinerias
    vector<bool> refinerias(cantPozos);
    for (int i = 0; i < cantPozos; ++i) {
        if(conexos.find_set(i) == i){
            refinerias[i] = true;
            costoTotal += costoRefineria;
            cantRef += 1;
        }
    }
//cout pedido
    cout << costoTotal << " " << cantRef << " " << conexionesMinimas.size() << endl;
    for (int i = 0; i < refinerias.size(); ++i) {
        if(refinerias[i])
            cout << i+1 << " ";
        }
    cout << endl;
    for (int i = 0; i < conexionesMinimas.size(); ++i) {
        cout << conexionesMinimas[i].pozoA+1 << " " << conexionesMinimas[i].pozoB+1 << endl;
    }
    return costoTotal;
}
```

```
vector<eje> generarArbolesMinimos(vector<eje>& ejes, int costoRefineria, int cantPozos){
    UnionFind bosqueMinimo(cantPozos);
    vector<eje> res;
//ordenamos los ejes segun su costo para obtener en tiempo constante, los menores
    sort(ejes.begin(), ejes.end());
    for (int i = 0; i < ejes.size(); ++i) {
//si agregar el eje no forma ciclo
    if(!bosqueMinimo.is_in(ejes[i].pozoA,ejes[i].pozoB)){
//lo uno y si cuesta menos que poner una refineria, lo agrego a res
        bosqueMinimo.union_set(ejes[i].pozoA, ejes[i].pozoB);
            if(ejes[i].costoTuberia < costoRefineria){
                eje conex;
                conex.pozoA = ejes[i].pozoA;
                conex.pozoB = ejes[i].pozoB;
                conex.costoTuberia = ejes[i].costoTuberia;
                res.push_back(conex);
            }
        }
    }
    return res;
}
```

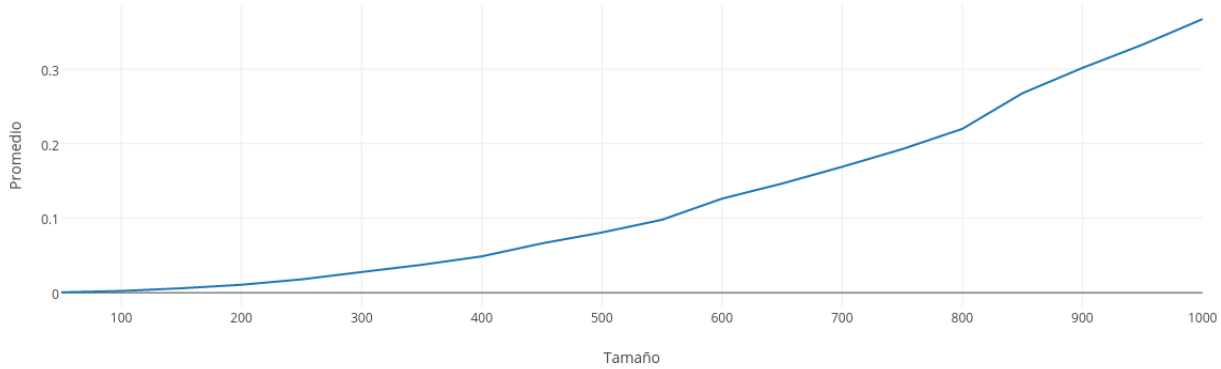
3.5. Experimentación

3.5.1. Constrastación Empírica de la complejidad

Para llevar a cabo esta experimentación, consideramos el peor caso posible de cantidad de ejes del grafo, es decir que habrá exactamente $n.(n-1)$ ejes (cada nodo puede conectarse con cualquier nodo del grafo, excepto consigo mismo), variando la cantidad de nodos.

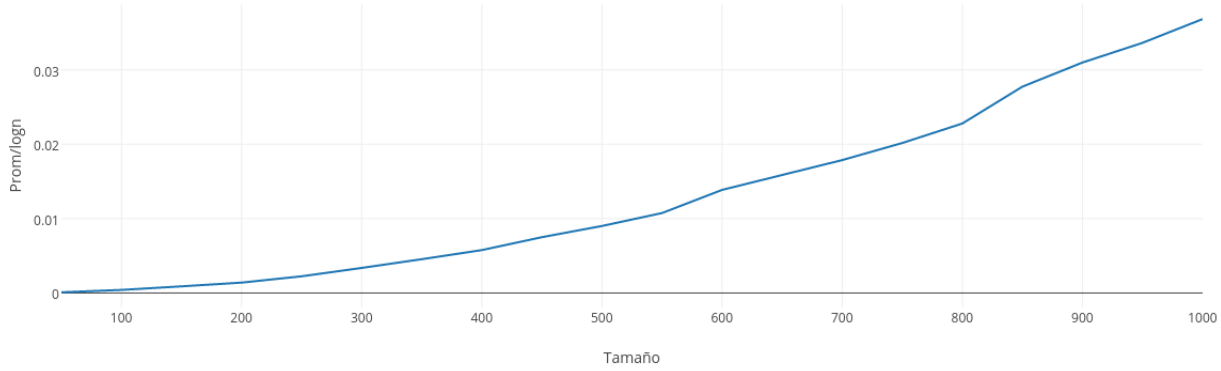
Los tiempos de ejecución para cada n (cantidad de nodos) fueron los siguientes:

n	Tiempo en segundos
50	0.0005254864
100	0.002332719
150	0.0059328642
200	0.01065945
250	0.0178728144
300	0.0277215188
350	0.0373807766
400	0.0487278266
450	0.0661190683
500	0.0807527896
550	0.0978193478
600	0.126140013
650	0.146378511
700	0.168847877
750	0.192539493
800	0.219705288
850	0.267457097
900	0.301437881
950	0.332602623
1000	0.366929358

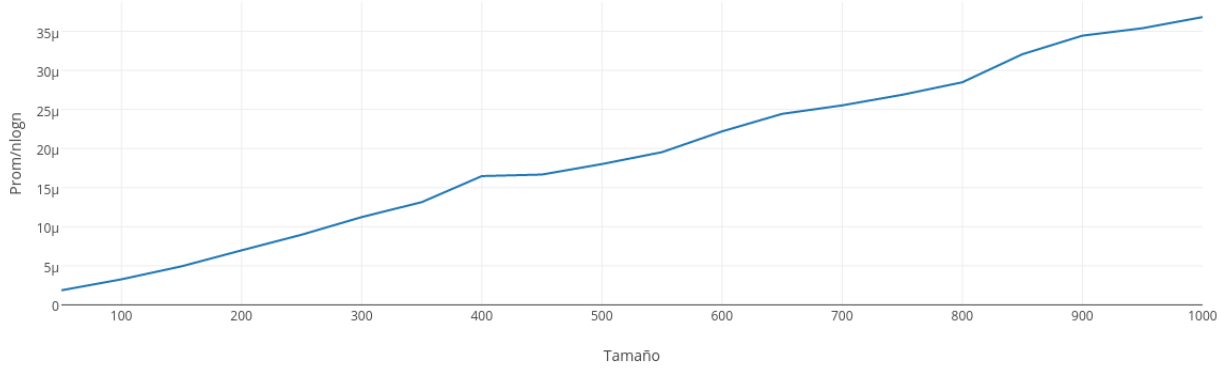


Dado que la Cota de Complejidad planteada teóricamente es de $O(n^2.log(n))$, era esperable que la curva sea una parábola creciente.

A simple vista, no se puede apreciar si la relación que tienen respecto de tamaño/tiempo es efectivamente la que buscamos (pues casi todas las curvas polinomiales tienen gráficos similares). Por este motivo, como siguiente paso decidimos comenzar a linealizar los tiempos, dividiendo a cada uno por $log(n)$.



La morfología de este gráfico es similar a la anterior, sigue siendo una parábola creciente, por lo tanto terminaremos de linealizar, dividiendo a cada uno por n, para ver si se trata de la parábola cuadrática.



Efectivamente puede observarse que el comportamiento es lineal. Por lo tanto, podemos afirmar que nuestra experimentación condice a la Cota Teórica planteada de $O(n^2.log(n))$.