

OBSERVACIONES DEL LA PRACTICA

DIEGO ALEJANDRO GONZÁLEZ VARGAS Cod 202110240

SEBASTIAN GUERRERO RIOS Cod 202021249

Preguntas de análisis

1) ¿Qué relación encuentra entre el número de elementos en el árbol y la altura del árbol?

RTA/ Se considera, de acuerdo con los resultados de la consola respecto a la altura del árbol y el número de elementos del mismo, que la relación que existe entre ellos es nula. Sin embargo, se hace necesaria una explicación para esta respuesta.

```
*****
Seleccione una opción para continuar
>2

Cargando información de crímenes ....
Crímenes cargados: 319073
Altura del arbol: 29
Elementos en el arbol: 1177
Menor Llave: 2015-06-15
Mayor Llave: 2018-09-03
*****
```

En este sentido, se tuvo en cuenta que el concepto de altura hace referencia a la cantidad máxima de saltos que se puede dar entre los diferentes niveles. Así las cosas, se puede decir que el árbol binario tendrá una misma cantidad de niveles que la altura sumada a 1. Así mismo, se reconoce también que la capacidad máxima de elementos en un árbol hace referencia a la n -ésima potencia de 2, siendo n el número del último nivel del árbol, que como se mencionó, es la misma altura. En consecuencia, consideramos que la cantidad de elementos podría tener alguna relación con la cantidad de niveles o la altura en la medida en que se correspondiera con la capacidad del mismo, pero bajo este esquema, la capacidad del árbol sería la siguiente:

$$2^n \text{elementos} \rightarrow 2^{29} \text{elementos} \rightarrow 536'870.912 \text{ elementos}$$

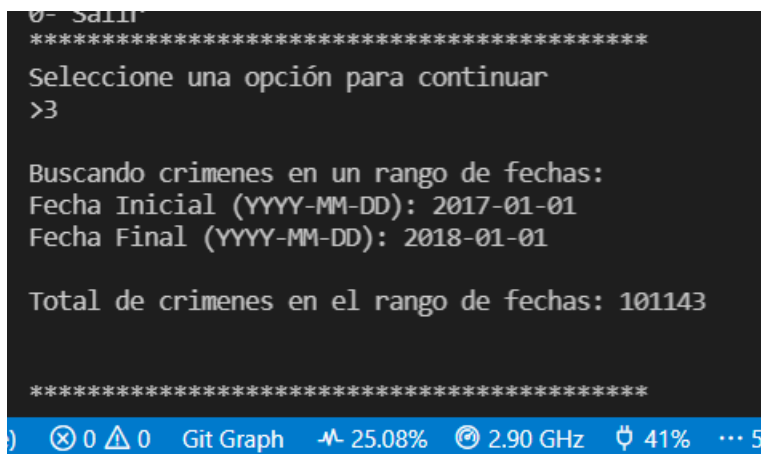
N: Número de la altura

Como se puede ver, no se acerca al valor de elementos recogido de manera práctica. Para poder verlo más claramente, es pertinente sacar un porcentaje de la cantidad de elementos en el BST con su capacidad máxima en un árbol completamente balanceado con 29 niveles:

$$\frac{1177 \text{ elementos reales}}{536'870.912 \text{ (capacidad)}} = 0,000002\%$$

Con este análisis podemos entender que este al ser un árbol binario organizado tiene una probabilidad grande de que este desbalanceado, esto significa que varios nodos pueden tener solo uno o cero hijos en varios niveles del árbol. En este caso, el árbol esta bastante desbalanceado y requiere muchos mas niveles para poder incluir todos los elementos <llave, valor>. En este árbol se evidencia que la relación entre niveles y elementos no es lineal, pero se evidencia que los niveles si incrementa a medida que se añaden elementos al árbol. Esto puede ser debido al elemento inicial encontrado y el orden en que entran.

- 2) ¿Si tuviera que responder esa misma consulta y la información estuviera en tablas de hash y no en un BST, cree que el tiempo de respuesta sería mayor o menor? ¿Por qué?



```
0- Salir
*****
Seleccione una opción para continuar
>3

Buscando crímenes en un rango de fechas:
Fecha Inicial (YYYY-MM-DD): 2017-01-01
Fecha Final (YYYY-MM-DD): 2018-01-01

Total de crímenes en el rango de fechas: 101143

*****
) 0 0 Git Graph 25.08% 2.90 GHz 41% ... 5.
```

RTA/ Se considera que para el caso de este enunciado se puede asegurar que el tiempo de respuesta sería mayor para el caso de una tabla de hash. En este sentido, se considera la cantidad de búsquedas dentro de la tabla de símbolos como un elemento clave para dar esta respuesta. Así pues, para el caso del BST utilizado en el código del laboratorio, al ser un mapa ordenado, se puede trabajar con el concepto de límites para organizar la respuesta que se le va a mostrar al usuario. En consecuencia, se pueden hacer únicamente 2 búsquedas en el BST para determinar la ubicación de los límites de nuestra respuesta de acuerdo con la solicitud de fechas del usuario. Posteriormente, se puede simplemente hacer un conteo de los elementos que halla entre estos límites y sus crímenes correspondientes.

Por otra parte, para el caso de una tabla de hash, se hace necesaria un ciclo de n de búsquedas, donde n es la cantidad de fechas diferentes que haya en el rango determinado por el usuario. Sin embargo, aunque este proceso puede no parecer tan diferente al anterior (teniendo en cuenta que para la implementación se crea un TAD lista con la respuesta y se va copiando cada dato a dicha lista), para el caso de la tabla de hash se van a realizar un número indeterminado de búsquedas infructuosas, de todas las fechas que estén en el rango y que no tengan elementos en la tabla de hash, lo que no sucede con el BST.

- 3) ¿Qué operación del TAD se utiliza para retornar una lista con la información encontrada en un rango de fechas?

```
*****
Seleccione una opción para continuar
>4

Buscando crímenes x grupo de ofensa en una fecha:
Fecha (YYYY-MM-DD): 2017-01-01
Ofensa: Vandalism

Total de ofensas tipo: Vandalism en esa fecha: 12

*****
```

RTA/ Se considera que para dar respuesta a esta pregunta es pertinente la utilización del código de la función que cumple con este requerimiento dentro de la práctica de laboratorio. Esta función es la de la opción 3, que efectivamente consulta crímenes en un rango de fechas, y que maneja el siguiente código:

```
182 def getCrimesByRange(analyzer, initialDate, finalDate):
183     """
184     Retorna el numero de crímenes en un rango de fechas.
185     """
186     lst = om.values(analyzer['dateIndex'], initialDate, finalDate)
187     totcrimes = 0
188     for lstdate in lt.iterator(lst):
189         totcrimes += lt.size(lstdate['lstcrimes'])
190     return totcrimes
191
```

Como se puede observar, la función del TAD orderedmap que se utiliza es la de values(), que permite a partir de 3 parámetros (el BST, límite inferior, límite superior) obtener una lista con toda la información de los valores de las llaves que se encuentran en el rango ocupado por estas 2 fechas estipuladas.