DEPARTAMENTO
INGENIERÍA INFORMÁTICA
Y CIENCIAS DE LA COMPUTACIÓN
FACULTAD DE INGENIERÍA UNIVERSIDAD DE CONCEPCIÓN

Estructuras de Datos y Algoritmos Avanzados (2022-2)

Proyecto 1: Árboles de búsqueda

Profesor: José Fuentes Sepúlveda
Ayudante: Alexander Irribarra Cortés

Objetivos

Los objetivos de este proyecto son:

- Mejorar la programación, compilación y ejecución de programas escritos en lenguaje *C++* u otros.
- Estudiar e implementar distintas variantes de árboles de búsqueda binarios.
- Analizar experimental y teóricamente el desempeño de las estructuras de datos implementadas.

1. Descripción del problema

Los árboles binarios de búsqueda son estructuras de datos que permiten, entre otras cosas, implementar diccionarios. Si bien para un mismo conjunto de datos existen múltiples árboles binarios de búsqueda válidos, los cuales se diferencian en su topología, algunos son más deseables que otros, pues reducen la altura del árbol, lo que conlleva a tener operaciones más rápidas.

En este proyecto se implementarán diversas variantes de árboles binarios de búsqueda auto-balanceables, los cuales utilizan distintas estrategias para acotar la altura del árbol. Entre las variantes a implementar se encuentras los árboles AVL, Red-Black y Splay Trees. Cada una de estas tiene ventajas y desventajas, por lo que en ciertos escenarios puede ser más útil utilizar una que las otras. Adicionalmente, se implementará como *baseline* un árbol binario de búsqueda sin balanceo.

Por simplicidad, los árboles a implementar solamente deben almacenar claves, y no valores. Es decir, su funcionalidad es equivalente a la de un *set*.

2. Objetivos específicos

La entrega del proyecto, que consiste de un informe y el código fuente, debe satisfacer los siguientes objetivos:

- Cada estructura de datos debe proveer las siguientes funcionalidades:
  - `void insert(int)`: Insertar una clave al *set*.
  - `bool search(int)`: Retorna *true* si y solo si la clave buscada se encuentra dentro del set *set*.
- Describir brevemente las características principales de los tipos de árboles implementados.
- Describir las decisiones de implementación más importantes.
- Plantear varias hipótesis sobre el rendimiento de las estructuras a comparar en escenarios específicos. Por ejemplo, "La estructura X es más rápida en consultas que la estructura Y cuando los datos cumplen cierta característica C". Se deben plantear (por lo menos) tantas hipótesis como integrantes tenga el equipo.
- Diseñar un experimento que permita verificar cada hipótesis.
- Ejecutar los experimentos y discutir los resultados obtenidos.

3. Condiciones

- El proyecto se realizará en grupos de dos o tres estudiantes. El informe debe reflejar claramente los autores del proyecto.
- Se deben implementar tantas variantes auto-balanceables de árboles como miembros tenga el grupo, además del *baseline*. Si el grupo es de dos personas, deben implementar el *baseline*, *AVL* y *Red-Black trees*. Si el grupo es de tres personas, además deben implementar *Splay trees*.
- Todas las implementaciones deben realizarse en el mismo lenguaje de programación y utilizar los mismos tipos de optimizaciones.
- Las implementaciones se pueden realizar en los lenguajes *C++*, *Java* o *Python*.
- Todas las implementaciones deben heredar del árbol binario de búsqueda sin balanceo.
- El *baseline* y el código de la evaluación experimental deben ser implementados entre todos los miembros del equipo, al igual que la definición de hipótesis y la redacción del informe.

- En el informe se debe describir el entorno de *hardware* y *software* en que se ejecutan los experimentos. Además, debe quedar clara la manera en que se generan u obtienen tanto los datos de entrada como los resultados reportados. Una persona externa con la información descrita en el informe, debiera ser capaz de replicar los experimentos y obtener las mismas conclusiones.
- Con el objetivo de reducir ruido en los resultados, las mediciones realizadas deben contemplar al menos 30 repeticiones, de las cuales se reportará la media y varianza.

4. Evaluación

El proyecto se realizará en grupos de dos o tres personas. Se debe subir a Canvas lo siguiente:

1. Un informe que:
   a) Incluya portada, descripción de la tarea, descripción de las soluciones propuestas, detalles de implementación, análisis teórico y análisis experimental, considerando las condiciones previamente estipuladas.
   b) Sea claro y esté bien escrito. Un informe difícil de entender será mal evaluado, aunque todo esté bien implementado. Quien revise el documento debe poder entender su solución solo mirando el informe.
   c) Esté en formato pdf.
2. Un archivo comprimido con todos los ficheros fuente implementados para solucionar la tarea. El informe debe hacer referencia a ellos y explicar en qué consiste cada uno.

Fecha de entrega: jueves 15 de septiembre de 2022 11:59PM