

LABORATORIO No. 7: Tablas de Símbolos Ordenadas

Objetivos

Comprender la implementación del Tipo Abstracto de Datos Tabla de Símbolos Ordenadas (Tree) y su uso para la solución de problemas.

Al finalizar este laboratorio el estudiante estará en capacidad de:

- Identificar las operaciones principales implementadas en el TAD Tree.
- Utilizar la estructura de datos Tabla de Símbolos Ordenadas para almacenar datos.
- Entender a usar un árbol Binario de Búsqueda (BST) como estructura de datos y su impacto en los órdenes de crecimiento temporal.
- Proponer la forma de utilización de un Árbol Binario de Búsqueda (BST) como alternativa de solución a un problema que implican el manejo de llaves ordenadas.
- Integrar los Árboles Binarios de Búsqueda (BSTs) con las otras estructuras de datos vistas en el curso.

Fecha Límite de Entrega

Sábado **16 de octubre** antes de la media noche (11:59 p.m.).

Preparación del Laboratorio

- Revisar el API del TAD Map ubicado en `DISClib\ADT\orderedmap.py`
- Revisar las estructuras de datos `orderedmapstructure.py` y `bst.py` ubicado en `DISClib\DataStructures`

Trabajo Propuesto

PASO 1: Copiar el ejemplo en su organización

Copie/Haga **Fork** del repositorio del laboratorio en su organización con el procedimiento aprendido en las prácticas anteriores.

El repositorio del proyecto base que utiliza este laboratorio es el siguiente:

- <https://github.com/ISIS1225DEVS/ISIS1225-SampleTree.git>

Antes de clonar el repositorio en su computador diríjase a su organización (Ej.: **EDA2021-1-SEC02-G01** para el primer grupo de la sección 2 del curso) y cambie el nombre del repositorio de acuerdo con el esquema **LabTrees-S<<XX>>-G<<YY>>** donde **XX** es el número de la semana de la práctica y donde **YY** es el número del grupo de trabajo. (Ej.: **LabTrees-S07-G01** para este **octavo laboratorio** hecho por el **grupo 1** de la **sección 2**).

Recuerde que **NO necesita** agregar la sección o el semestre en este nombre porque ya está identificado en su organización.

PASO 2: Descargar el ejemplo

Después de renombrar el proyecto dentro de su organización ya puede clonar el proyecto. Descargue el código en su máquina local (**git clone**) siguiendo lo aprendido en las practicas anteriores.

Recuerde modificar el **README** del repositorio para incluir los nombres de los integrantes del grupo.

PASO 3: Ejecutar y explorar el ejemplo

El proyecto **SampleTree** busca familiarizarlos con el TAD Tabla de Símbolos Ordenadas (Tree) y su uso para solucionar problemas y una forma de probar su desempeño en aplicaciones MVC.

Antes de iniciar a explorar y modificar el ejemplo, recuerde descargar los datos de trabajo **Boston Crimes** disponibles en el portal oficial del curso en BrightSpace. Descargue el **Zip**, descomprímalo y guarde los archivos CSV en la carpeta ***/Data/** de su copia local de código.

Diríjase al archivo **view.py** y ejecútelo, y seleccione secuencialmente la **opción 1** y **2** para iniciar el analizador y cargar información respectivamente.

```
*****
Bienvenido
1- Inicializar Analizador
2- Cargar información de crímenes
3- Consultar crímenes en un rango de fechas
4- Consultar crímenes por código y fecha
0- Salir
*****
Seleccione una opción para continuar
>2

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

Al cargar la información, verá el número de registros cargados, la altura del árbol (BST), el número de nodos en el árbol (BST), la menor llave encontrada, y la mayor llave encontrada. El árbol guarda un índice con parejas llave-valor, donde la llave es la fecha del crimen (YYYY-MM-DD) y el valor, los crímenes ocurridos ese día.

A continuación, responda la siguiente pregunta y registre su respuesta en el documento de observaciones:

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

Ahora, ejecute la **opción 3**, que retorna todos los crímenes encontrados en un rango de fechas. Utilice fechas de inicio y final en formato YYYY-MM-DD y dentro de las fechas del archivo (2015 a 2018). Por ejemplo, como fecha de inicio puede dar: 2017-01-01 y como fecha final 2018-01-01

```
*****
Bienvenido
1- Inicializar Analizador
2- Cargar información de crímenes
3- Consultar crímenes en un rango de fechas
4- Consultar crímenes por código y fecha
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
```

Verá como resultado que en ese periodo de tiempo se produjo un total de 101143 crímenes.

Responda la siguiente pregunta y registre su respuesta en el documento de observaciones:

b) ¿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?

A continuación, ejecuten la **opción 4**, que retorna todos los crímenes de un tipo determinado y en una fecha específica. Por ejemplo, como fecha de puede dar: 2017-01-01 y como tipo de crimen: "Vandalism". Deberá ver como resultado el total de crímenes en esa fecha y de ese tipo.

```
*****
Bienvenido
1- Inicializar Analizador
2- Cargar información de crímenes
3- Consultar crímenes en un rango de fechas
4- Consultar crímenes por código y fecha
0- Salir
*****
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
```

Responda la siguiente pregunta y registre su respuesta en el documento de observaciones:

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

PASO 4: Estudiar el ejemplo en VSCode

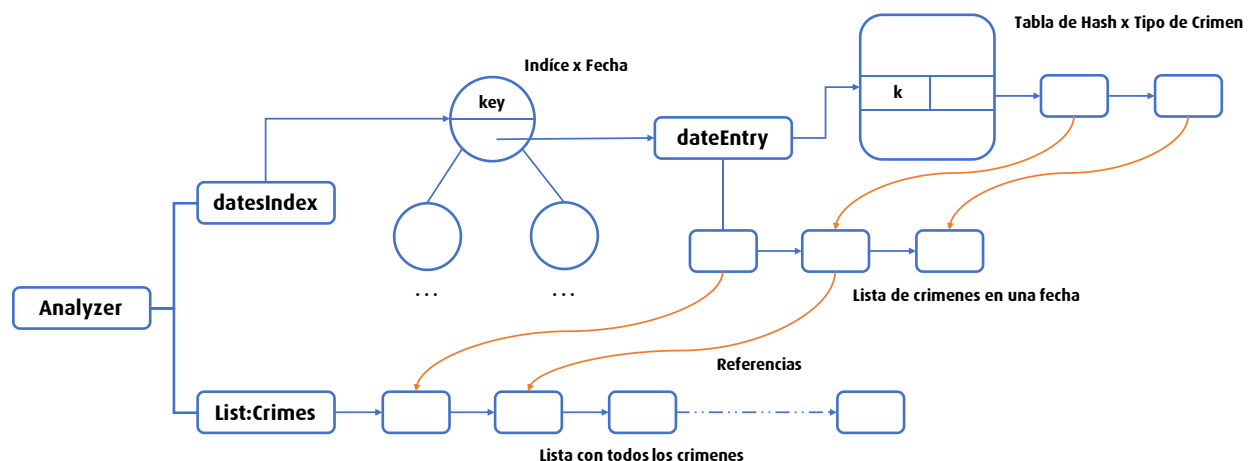
En el archivo **model.py** del ejemplo inspeccione el código para entender los TADs y estructuras de datos utilizados.

Lo primero que deben estudiar, es cómo está representado un analizador de crímenes. En el ejemplo el analizador tiene dos elementos: Una lista con todos los crímenes cargados y un índice (**dateIndex**) representado con un mapa ordenado, que en este caso es un BST.

Cada nodo del árbol tiene una pareja llave-valor. La llave es una fecha y el valor, a su vez, es de nuevo una pareja de elementos: a) La lista de todos los crímenes ocurridos en dicha fecha y una tabla de hash, con una lista de los crímenes en esa fecha, agrupados por tipo de crimen (la llave para la tabla de hash es el tipo de crimen).

Revisen el código hasta entender cómo se construye el árbol y cómo se usan el árbol y la tabla de hash para solucionar el problema.

La siguiente imagen ilustra las estructuras de datos utilizadas.



PASO 5: Actualizar los repositorios

Para el repositorio del laboratorio confirme los cambios con los comandos **Commit** y **Push** en la rama **main** local y de GitHub con el comentario *"laboratorio 7 - Entrega final"* antes de la fecha límite de entrega.

PASO 6: Revisar entregables de la practica

Finalmente, para realizar la entrega del laboratorio revise que sus entregables de la practica estén completos. Para ello, siga las siguientes indicaciones:

- 1) Acceso al profesor de laboratorio y los monitores de su sección a la organización del grupo.
- 2) **README** del repositorio con los datos completos de los integrantes del grupo (nombre completo, correo Uniandes y código de estudiante).

- 3) Enlace al repositorio GitHub **LabTrees-S<<XX>>-G<<YY>>** con rama **Main** actualizada con el comentario "*Laboratorio 7 – Entrega final*" antes del límite de entrega.
- 4) Incluir en repositorio del laboratorio en la carpeta **Docs** el documento **observaciones-lab8.pdf** con las respuestas a las preguntas de observación.
 - a) ¿Qué relación encuentra entre el número de elementos en el árbol y la altura del árbol?
 - b) ¿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?
 - c) ¿Qué operación del TAD se utiliza para retornar una lista con la información encontrada en un rango de fechas?

PASO 7: Compartir resultados con los evaluadores

Envíe el **enlace (URL)** del repositorio por **BrightSpace** antes de la fecha límite de entrega.

Recuerden que cualquier documento solicitado durante la práctica debe incluirse dentro del repositorio GIT y solo se calificarán los entregables hasta el último **COMMIT** realizado previo a la media noche (11:59 PM) del 16 de octubre de 2021.