

Estructuras de Datos

Proyecto Segunda Evaluación – PAO 1 2023

Diccionario de palabras con árbol Trie

LEA ESTE DOCUMENTO DETENIDAMENTE

Este es un proyecto **grupal**. Por tanto, todos los miembros del grupo deben involucrarse en el diseño e implementación de la solución que se solicita. La entrega del proyecto incluirá un reporte de auto y coevaluación, donde cada miembro del grupo calificará sus contribuciones al proyecto y las de sus compañeros.

Es su responsabilidad agruparse a tiempo para la entrega del proyecto. Usted debe indicar su grupo en el Aula Virtual hasta el miércoles 7 de junio. Los estudiantes que no se hayan unido a ningún grupo para entonces, serán asignados aleatoriamente por el profesor. A partir de la fecha indicada, no se aceptarán cambios en la conformación de los grupos. **EN NINGUNA CIRCUNSTANCIA SE ACEPTARÁN ENTREGAS INDIVIDUALES.**

Si algún grupo experimenta el incumplimiento sistemático de alguno de sus miembros, esta situación debe ser reportada al profesor **TAN PRONTO COMO SE PRESENTE**. Cualquier estudiante cuyo incumplimiento sea comprobado, será retirado del grupo y **DEBERÁ IMPLEMENTAR EL PROYECTO DE MANERA INDIVIDUAL SOBRE EL 50% DEL PUNTAJE POSIBLE.**

Si algún estudiante siente que otro (u otros) toma(n) el control del grupo y no permite(n) a los demás miembros hacer aportes o colaborar en el proyecto, debe reportar esta situación al profesor **TAN PRONTO COMO SE PRESENTE**. Cualquier estudiante que pretenda monopolizar el trabajo del proyecto o impida a los demás hacer aportes será separado del grupo y **DEBERÁ IMPLEMENTAR EL PROYECTO DE MANERA INDIVIDUAL Y SOBRE EL 50% DEL PUNTAJE POSIBLE.**

Si algún estudiante reporta de manera extemporánea cualquiera de las dos situaciones detalladas en los dos párrafos precedentes, será penalizado con su 25% de la nota total obtenida en el proyecto. Reporte novedades al profesor **TAN PRONTO COMO ÉSTAS SE PRESENTEN.**

Considere las políticas de buena conducta académica que se explicaron en la primera clase del curso respecto al plagio y demás violaciones del código de Ética de la ESPOL. Los autores del proyecto deben ser usted y sus compañeros de grupo. Cualquier indicio de lo contrario, será reportado a las unidades correspondientes para el tratamiento pertinente.

Finalmente, **siga las instrucciones de este documento** para evitar inconvenientes. Si tiene dudas, consulte al profesor en lugar de asumir cosas que pueden ser incorrectas.

El incumplimiento de instrucciones explícitas indicadas en este documento o indicadas en nuestras sesiones teóricas derivará, de manera inapelable, en la penalización y rebaja de puntos sobre su entrega.

Introducción

El objetivo de este proyecto es que usted y sus compañeros de grupo tengan la oportunidad de aplicar los conceptos y habilidades adquiridos en el curso de Estructuras de Datos en un contexto práctico. Más específicamente, usted desarrollará una aplicación que permitirá a los usuarios interactuar con un diccionario de palabras utilizando una estructura de árbol llamada Trie. Como parte de este proceso, usted deberá:

- Diseñar e implementar tipos de datos para representar las entidades involucradas en un administrador de un diccionario de palabras.
- Implementar árboles para organizar los datos manejados en la aplicación.
- Implementar algoritmos que permitan manipular y acceder a los datos de la aplicación mediante estructuras lineales y no lineales.
- Diseñar e implementar una interfaz gráfica que permita interactuar con los datos del diccionario.

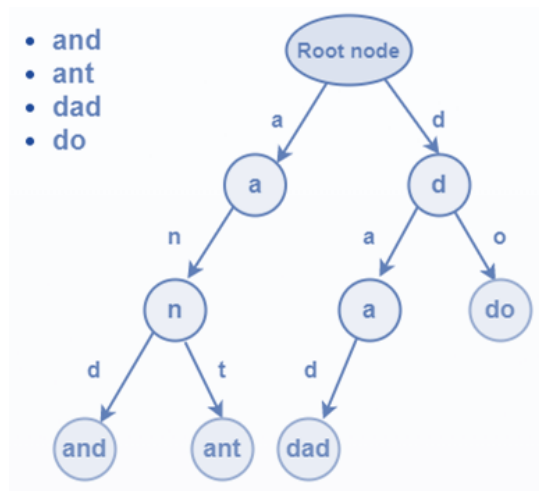
El Árbol Trie

En este proyecto, usted representará al diccionario de la aplicación mediante un árbol multicaminos llamado Trie (<https://es.wikipedia.org/wiki/Trie>).

Un árbol Trie (también conocido como árbol prefijo) es una estructura de datos jerárquica utilizada principalmente para almacenar y buscar palabras o cadenas de caracteres. El nombre "Trie" proviene del término inglés "*retrieve*" (que significa recuperar), ya que su diseño está optimizado para recuperar y buscar rápidamente palabras o prefijos comunes.

En un árbol Trie, cada nodo representa una letra y cada camino desde la raíz hasta un nodo hoja forma una palabra. Las palabras se almacenan de manera que los caracteres comunes en la parte superior del árbol se comparten entre múltiples palabras, lo que reduce significativamente el espacio de almacenamiento requerido.

A continuación, se muestra un árbol Trie con las palabras "and", "ant", "dad", y "do":



Características principales de un árbol Trie:

Raíz: El nodo superior del árbol Trie, que no representa ningún carácter en particular, pero tiene enlaces a los nodos que representan los posibles caracteres iniciales de las palabras.

Nodos y enlaces: Cada nodo en el árbol Trie representa una letra y tiene enlaces a los nodos que representan las letras siguientes en la palabra.

Nodos hoja: Los nodos hoja representan el final de una palabra y no tienen enlaces adicionales.

Requerimientos Mínimos

La interfaz gráfica de su aplicación deberá permitir:

1. **Inserción de palabras:** La interfaz gráfica debe permitir a los usuarios agregar nuevas palabras al diccionario. Al ingresar una nueva palabra, la aplicación la insertará en el árbol Trie asegurándose de mantener la estructura y optimizando el espacio.
2. **Búsqueda de palabras:** Los usuarios podrán buscar palabras en el diccionario utilizando la interfaz gráfica. Al ingresar una palabra en un cuadro de búsqueda, la aplicación recorrerá el árbol Trie para verificar si la palabra existe en el diccionario. Su aplicación debe informar apropiadamente al usuario sobre los resultados de la búsqueda.
3. **Completar palabras:** Cuando los usuarios estén escribiendo una palabra en el cuadro de búsqueda, la interfaz gráfica puede proporcionar sugerencias de palabras completas que coincidan con las letras ingresadas hasta el momento. Esto se logra recorriendo el árbol Trie hasta la ubicación donde coinciden las letras ingresadas y luego buscando todas las palabras que siguen esa ruta.
4. **Eliminación de palabras:** Los usuarios pueden tener la opción de eliminar palabras del diccionario a través de la interfaz gráfica. Al eliminar una palabra, la aplicación ajustará el árbol Trie adecuadamente para mantener la consistencia de la estructura.
5. **Carga y guardado de diccionario:** La interfaz gráfica debe permitir a los usuarios cargar y guardar el diccionario desde y hacia un archivo de texto, donde cada palabra debe aparecer en una fila. Esto facilitará la manipulación de un diccionario existente y la reutilización de palabras agregadas previamente. Las palabras insertadas por el usuario deben almacenarse en el archivo y estar disponibles la próxima vez que se ejecute su aplicación.

Sobre el Diseño de su Interfaz Gráfica

Usted y sus compañeros de grupo son responsables de decidir el diseño y la apariencia de su interfaz gráfica. Aunque este no es un curso de Interacción Humano-computador, se recomienda que su aplicación sea estéticamente agradable, intuitiva, y fácil de usar. Sea cuidadoso en el uso de colores y la selección y organización de los elementos de su interfaz gráfica.

A diferencia del proyecto parcial, la aplicación que se solicita en este documento es bastante simple. Esencialmente, la interfaz gráfica debe constar de un cuadro de texto donde los usuarios escriben palabras para buscar/insertar/eliminar en/del diccionario. Su aplicación debe, además, incluir varios botones (botón de búsqueda, botón de inserción, botón de eliminación, botones para guardar y cargar el diccionario, etc.) y debe contar con una sección donde se muestren los resultados de la búsqueda y las palabras sugeridas a medida que el usuario escribe.

Su interfaz gráfica debe ser clara y concisa, y debe mostrar mensajes informativos al usuario sobre el resultado de las operaciones realizadas. También se puede considerar la inclusión de elementos adicionales (ver sección sobre Funcionalidad Opcional), como botones para mostrar estadísticas del árbol, etc.

Funcionalidad Opcional

Las funcionalidades descritas arriba son obligatorias y constituyen los requerimientos mínimos para que su proyecto sea aceptado y admitido a ser calificado **sobre hasta el 80%** de la nota máxima. Este proyecto, sin embargo, es también una oportunidad para ser creativo. Así, usted y sus compañeros de grupo pueden implementar funcionalidad extra que contribuya a una mejor versión de la aplicación solicitada.

A continuación, se listan algunas funcionalidades opcionales que usted y sus compañeros de grupo podrían implementar para aplicar a puntos adicionales:

- **Panel de estadísticas:** Muestra algunas estadísticas sobre los datos almacenados en el diccionario de palabras. Por ejemplo, la cantidad total de palabras almacenadas, la cantidad de palabras por cada letra del alfabeto, un histograma de longitudes, etc.
- **Búsqueda por prefijo:** Agregar una función que permita a los usuarios buscar todas las palabras que comiencen con un prefijo específico. Por ejemplo, al ingresar "in", se mostrarían palabras como "iniciar", "inspección", etc.
- **Búsqueda inversa:** Implementar una búsqueda que permita a los usuarios buscar palabras que terminen con una determinada terminación. Por ejemplo, al ingresar "ción", se mostrarían palabras como "nación", "información", etc.
- **Búsqueda aproximada:** Agregar una búsqueda que encuentre palabras similares a la consultada, utilizando algoritmos de edición (como la distancia de Levenshtein) para encontrar palabras con solo un pequeño número de caracteres diferentes. Las palabras "cat" y "car" son aproximadas, pues ambas difieren únicamente en la última letra.
- **Modo de juego:** Crear un juego basado en la búsqueda y manipulación de palabras en el Trie. Por ejemplo, un juego de "buscar palabras" donde el usuario tiene que encontrar todas las palabras que se pueden formar a partir de una combinación de letras.

En general, si su implementación es excepcional, su proyecto podría superar el 100% de la nota. Usted podría utilizar los puntos excedentes para compensar otros rubros de la materia (p. ej., lecciones o tareas).

Modelamiento del Problema y Consideraciones Técnicas

Como habrá notado, parte de la funcionalidad solicitada arriba puede ser modelada con árboles multicaminos de Strings. Dados los objetivos que se persiguen en este curso, usted y sus compañeros de grupo serán responsables de implementar todas las estructuras no lineales que necesiten en la solución de este problema. A diferencia del proyecto de la primera evaluación, está **permitido** utilizar todas las estructuras lineales provistas por Java.

Su aplicación debe consistir en una interfaz gráfica de JavaFX que permita al usuario realizar las operaciones indicadas arriba y todas las demás que sean pertinentes para lograr los requerimientos mínimos.

Las palabras que componen el diccionario de la aplicación deben ser "recordadas" por la herramienta. Para esto, deben almacenarse en archivos de texto, de tal manera que los datos del sistema sean persistentes y estén disponibles cada vez que se ejecute su aplicación, **sin necesidad de que alguien tenga que ingresar toda la información desde cero.**

Usted es responsable de decidir la estrategia mediante la cual logrará la persistencia de los datos. Es decir, queda a discreción suya y de sus compañeros de grupo utilizar archivos de texto plano o serializar la información en archivos binarios. **No utilice almacenamiento que requiera motores de bases de datos.**

Entregables

El proyecto de Netbeans que implementa correctamente la interfaz gráfica final de su proyecto, con –al menos– las funcionalidades mínimas detalladas anteriormente, debe ser entregado hasta las **11:59 pm del jueves 24 de agosto de 2023**.

Su entrega deberá incluir además un archivo **.docx**. La plantilla específica de este documento se adjunta a este documento y deberá incluir:

- *Screenshots* de su interfaz explicando su funcionalidad,
- Una imagen de su clase principal, y
- Tabla de auto- y co-evaluación de los miembros del grupo

Su proyecto de Netbeans y el archivo **.docx** deben ser entregados a través del Aula Virtual en un único archivo comprimido **.zip**.

No se aceptarán entregas atrasadas, sin reporte, o que no implementen los requerimientos mínimos.

Recuerde que este **es un proyecto grupal**. Esto significa que usted debe trabajar **con** sus compañeros de grupo para sacar el proyecto adelante. Su grupo deberá reunirse periódicamente para coordinar y discutir acciones relacionadas al proyecto. Dividir funcionalidad y repartir responsabilidades podría no ser la mejor estrategia, sobre todo en las fases tempranas del proyecto.

Presentaciones

Usted y sus compañeros de grupo deberán presentar su proyecto (el código subido al Aula Virtual) al profesor en vivo, en una reunión que tendrá lugar de manera remota el **viernes 25 de agosto** en un horario que será anunciado con anticipación. Esta reunión tendrá una duración de **15 MINUTOS** por grupo, será grabada, y se efectuará únicamente entre el profesor y **TODOS** los miembros del grupo. Las observaciones sobre su entrega serán dadas de manera verbal en esta reunión.

Solo los grupos que hayan entregado un proyecto funcional podrán y deberán presentarse a esta reunión.

La reunión iniciará con uno de los miembros del grupo compartiendo su pantalla para mostrar su proyecto en funcionamiento. Esta demostración debe evidenciar que el grupo ha cumplido con los requerimientos mínimos del proyecto y debe mencionar explícitamente toda funcionalidad extra que hayan implementado. Su grupo debe definir con anticipación qué miembro del grupo estará a cargo de esta presentación inicial.

Las cámaras y micrófonos de todos los miembros del grupo deben permanecer encendidos mientras dure la reunión con el profesor.

La presentación inicial **NO** debe ser sobre los requerimientos del proyecto --- el profesor tiene claro lo que se ha solicitado. Mas bien, debe enfocarse en lo que su grupo ha implementado. El profesor también conoce los nombres de los integrantes del grupo, por tanto, no es necesario que usted presente a sus compañeros.

Enfóquese en mostrar su proyecto en ejecución y en comunicar al profesor el resultado que su grupo ha logrado.

Durante la reunión, cada integrante del grupo deberá responder, al menos, una pregunta sobre el proyecto. Para esto, en la presentación, todos los miembros del grupo deben tener el código subido al Aula Virtual abierto en sus computadoras y deben estar listos para compartir su pantalla al responder las preguntas del profesor. Sus respuestas serán calificadas y la calificación individual recibida será uno de los dos factores multiplicativos que contribuirán a su nota en el proyecto.

Los detalles del enlace mediante el cual se efectuará la reunión serán indicados oportunamente.

Instrucciones Finales

Recuerde: **Este es un proyecto grupal.**

Finalmente, **siga las instrucciones de este documento** para evitar inconvenientes. Si tiene dudas, consulte al profesor (en lugar de asumir cosas que pueden ser incorrectas). Los ayudantes de la materia podrían no ser la fuente de información más confiable cuando se trata de responder preguntas sobre el proyecto.

Sea oportuno(a) en cuanto a sus preguntas: **No se atenderán consultas sobre el proyecto después del 1 de agosto.** Tampoco se responderán preguntas que consulten sobre instrucciones indicadas explícitamente en este documento. Antes de hacer preguntas, **LEA ESTE DOCUMENTO DETENIDAMENTE.**

¡Muchos éxitos a todos!