

CS 201: Data Structures II

Fusion Trees

L5-Group-4

Spring 2023

1 Group Members

1. Ali Muhammad Asad aa07190
2. Iqra Ahmed ia07674
3. Mohit Rai mr06638
4. Saira Junaid sj07420
5. Anoosha Hasan ah05860

2 Data Structure

A Fusion Tree is a modification of a B-Tree that implements an associative array on w -bit integers in a known universe size. An associative array is an abstract data type that stores a collection of pairs such as key-value pair. Fusion Trees are mostly used when our universe size is large, while providing linear - $O(n)$ - space complexity, and $O(\log n)$ search time complexity making it faster than a traditional self-balancing tree.

3 Application

Fusion trees are used in various applications such as computational geometry, databases, and data compression, machine learning and web search engines. They can be used to efficiently solve geometric problems like nearest neighbor queries and range search in higher dimensions, efficiently index and search large databases of numerical data, building blocks for various machine learning algorithms such as decision trees, and build indexes for search engines that can efficiently handle queries that involve numerical data.

We will mainly be focusing on applying Fusion Trees to efficiently index and search large databases of numerical data or any of the other emerging fields. We might also try and compare its performance with other trees or self-balancing trees.

4 Functionality

For the Fusion Tree, we will be implementing the search, insert, and remove operation over our data set through the application. The application will have options to select whether a user wants to find an existing value, if the value does not exist then a prompt will be shown. The user can insert a value or update an existing value to the data set, or the user can remove a value existing within the data set.

5 Datasets

We will be working over real life data sets compiled using a csv file. The csv file will be read and a tree will be made using the data set.

6 Work Distribution

Fill in the table which indicates the work distribution of each member.

Item	Activity	ID
1	Implementing Fusion Tree	aa07190, ia07674, ah05860
2	Handling the Data Set	aa07190, ia07674, ah05860
3	Implementing the interface	mr06638, sj07420

7 Attribution

Please cite any sources that you use, especially if using AI.

References

- [1] OpenAI. *ChatGPT*. [Online]. Available: <https://openai.com/chat> . Last accessed: 3 April 2023.
Prompt: Fusion Tree Data Structure
- [2] *tutorialspoint*. [Online]. Available: <https://www.tutorialspoint.com/cplusplus-program-to-implement-fusion-tree> . Last accessed: 3 April 2023
- [3] *stackoverflow*. [Online]. Available: <https://stackoverflow.com/questions/3878320/understanding-fusion-trees>. Last Accessed: 3 April 2023