

# Final Project - Inverted Index and Comparative Analysis

Arthur Rabello Oliveira<sup>1</sup>, Gabrielle Mascarello, Eliane Moreira, Nicolas Spaniol, Gabriel Carneiro

## Abstract

We present the implementation and comparative analysis of three fundamental data structures for indexing and searching textual documents: the Binary Search Tree (BST), the AVL Tree, and the Red-Black Tree (RBT). Each structure was implemented with its core operations, including insertion and search. Unit tests were developed to validate the correctness and performance of these implementations. We also provide a further comprehensive comparative study of the three trees based on their time complexity, balancing efficiency, and suitability for document indexing. The results demonstrate the trade-offs between implementation complexity and query performance, offering insights into the practical considerations for choosing appropriate search tree structures in information retrieval systems.

## Contents

1. Introduction .....	3
1.1. Motivation .....	3
1.2. Problem Statement .....	3
1.3. Overview of Search Trees .....	3
2. Data Structures Overview .....	3
2.1. Binary Search Tree (BST) .....	3
2.2. AVL Tree .....	3
2.3. Red-Black Tree (RBT) .....	3
2.4. Inverted Index .....	3
2.5. Comparison of Properties .....	3
3. Implementations .....	3
3.1. Binary Search Tree (BST) .....	3
3.1.1. Algorithms .....	3
3.1.2. Complexity Analysis .....	3
3.2. AVL Tree .....	3
3.2.1. Algorithms .....	3
3.2.2. Complexity Analysis .....	3
3.3. Red-Black Tree (RBT) .....	3
3.3.1. Algorithms .....	3
3.3.2. Complexity Analysis .....	3
3.4. Inverted Index .....	3
3.4.1. Algorithms .....	3
3.4.2. Complexity Analysis .....	3
4. Testing and Validation .....	3
4.1. Unit Testing Method .....	3
4.1.1. Binary Search Tree (BST) .....	3
4.1.2. AVL Tree .....	3
4.1.3. Red-Black Tree (RBT) .....	3
5. Comparative Analysis .....	3

---

<sup>1</sup>Escola de Matemática Aplicada, Fundação Getúlio Vargas (FGV/EMAp), email: [arthur.oliveira.1@fgv.edu.br](mailto:arthur.oliveira.1@fgv.edu.br)

5.1. The Experiment .....	4
5.2. Memory Usage .....	4
5.3. Time Complexity .....	4
6. Conclusion .....	4
6.1. Summary of Findings .....	4
7. Source code .....	4

# **1. Introduction**

## **1.1. Motivation**

## **1.2. Problem Statement**

## **1.3. Overview of Search Trees**

# **2. Data Structures Overview**

## **2.1. Binary Search Tree (BST)**

## **2.2. AVL Tree**

## **2.3. Red-Black Tree (RBT)**

## **2.4. Inverted Index**

## **2.5. Comparison of Properties**

# **3. Implementations**

## **3.1. Binary Search Tree (BST)**

### **3.1.1. Algorithms**

### **3.1.2. Complexity Analysis**

## **3.2. AVL Tree**

### **3.2.1. Algorithms**

### **3.2.2. Complexity Analysis**

## **3.3. Red-Black Tree (RBT)**

### **3.3.1. Algorithms**

### **3.3.2. Complexity Analysis**

## **3.4. Inverted Index**

### **3.4.1. Algorithms**

### **3.4.2. Complexity Analysis**

# **4. Testing and Validation**

## **4.1. Unit Testing Method**

### **4.1.1. Binary Search Tree (BST)**

### **4.1.2. AVL Tree**

### **4.1.3. Red-Black Tree (RBT)**

# **5. Comparative Analysis**

## **5.1. The Experiment**

## **5.2. Memory Usage**

## **5.3. Time Complexity**

## **6. Conclusion**

### **6.1. Summary of Findings**

## **7. Source code**

(repository)