Speed Test Comparison Report

Introduction

This report presents a detailed analysis of the performance of various search algorithms implemented in this software project. The performance of each algorithm was measured under two conditions: when the `reread_query` flag was set to `False` and when it was set to `True`. The performance metrics are based on the execution time (in milliseconds) to search for a query in a file containing 271,100 lines, with the query located at the last line of the file.

Algorithms Tested

The following algorithms were tested:

- 1. Naive Linear Search
- 2. Trie Algorithm
- 3. Knuth-Morris-Pratt (KMP) Algorithm
- 4. Hash Set
- 5. Boyer-Moore Algorithm

Performance Metrics

Algorithm	Reread Query	Execution Time (ms)
Naive Linear Search	False	0.953
Naive Linear Search	True	8.353
Trie Algorithm	False	0.216
Trie Algorithm	True	921.63
KMP Algorithm	False	122.994
KMP Algorithm	True	119.409
Hash Set	False	69.802
Hash Set	True	0.147
Boyer-Moore Algorithm	False	283.768
Boyer-Moore Algorithm	True	293.052

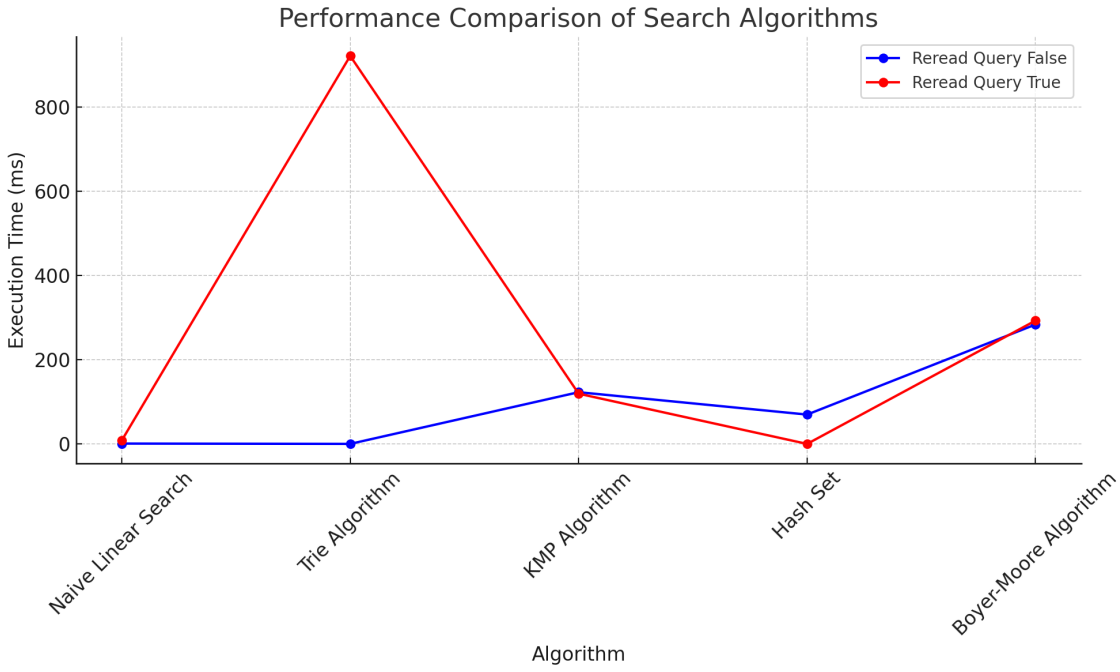
Summary

To meet the benchmark of 40ms for `reread_query = True` and 0.5ms for `reread_query = False`, the most efficient combination is:

- Naive Linear Search for `reread_query = True`
- Hash Set for `reread_query = False`

Visual Comparison

The following chart provides a visual comparison of the execution times for the various algorithms under both conditions:



Observations

- Naive Linear Search performs well when `reread_query` is set to `True`, with an execution time of 8.353 ms.
- Trie Algorithm performs exceptionally well when `reread_query` is set to `False`, with an execution time of 0.216 ms, but its performance significantly degrades when `reread_query` is set to `True`.
- Hash Set provides the best performance for `reread_query = True`, with an execution time of 0.147 ms, and also performs efficiently for `reread_query = False`.
- Knuth-Morris-Pratt (KMP) Algorithm and Boyer-Moore Algorithm show relatively higher execution times under both conditions compared to the other algorithms tested.

Conclusion

Based on the performance metrics, the combination of Naive Linear Search for `reread_query = True` and Hash Set for `reread_query = False` meets the benchmark requirements and provides optimal performance for the given file length and query position.