# Preliminary Analysis

|  |  |  |
| --- | --- | --- |
|  | Normal intersection algorithm | Heuristic-based on posting length |
| Time complexity | O(q∙N) | O(q∙N+q∙logq) |
| Space complexity | O(1) | O(logq) to O(q) |

* Where q is the number of query components
* Where N is the number of documents

# Experiments

## Using the top 10 most frequent words in the English language

### Hypothesis

And optimization will do worse than the normal intersection algorithm as it will do more work for the same result.

## Findings

Running the experiment on the park corpus we found the heuristic-based algorithm did twice as slow as the normal algorithm.

## Using the top 10 most frequent words in the English language alongside a word, not in the corpus

### Hypothesis

The heuristic-based algorithm will perform better than the normal intersection algorithm as it will get an empty list immediately and terminate

## Findings

The heuristic-based algorithm does in fact perform better than the other algorithm, achieving 0 ms query times

# Conclusion

The heuristic-based algorithm is not optimal as it can perform worst the than the other intersection algorithm. In other situations, it can perform better.