

Assignment –5

Explanation

Boyer-Moore Algorithm

The Boyer-Moore algorithm improves the efficiency of string matching by skipping sections of the text based on preprocessed information. It uses two main heuristics:

1. **Bad Character Heuristic:** When a mismatch occurs, the pattern is shifted to align the next occurrence of the mismatched character in the pattern with its occurrence in the text. If the mismatched character is not present in the pattern, the pattern is shifted past the mismatched character.
2. **Good Suffix Heuristic:** When a mismatch occurs, the pattern is shifted to align the last occurrence of a matching suffix with the suffix in the text.

Why Boyer-Moore Outperforms Others

3. **Skipping Sections:** The algorithm can skip sections of the text that do not need to be compared, reducing the number of comparisons significantly.
4. **Efficient Mismatches Handling:** Both heuristics allow for efficient handling of mismatches, often allowing the algorithm to skip multiple characters in the text.
5. **Best-Case Efficiency:** In the best-case scenario, the Boyer-Moore algorithm runs in $O(n/m)$ time, where n is the length of the text and m is the length of the pattern, which is highly efficient.

This implementation focuses on finding the last occurrence of the pattern in the text, leveraging the Boyer-Moore algorithm's strengths to provide efficient substring search functionality.